IMPERIAL INSTITUTE
OF
AGRICULTURAL RESEARCH, PUSA.

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A

CONTAINING PAPERS OF A MATHEMATICAL AND
PHYSICAL CHARACTER

VOL. CXXI

314736
IARI

LONDON
PRINTED FOR THE ROYAL SOCIETY AND SOLD BY
HARRISON AND SONS, LTD., ST. MARTIN'S LANE
PRINTERS IN ORDINARY TO HIS MAJESTY

DECEMBER 1928

LONDON
HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY,
ST. MARTIN'S LANE.

CONTENTS.

SERIES A. VOL. CXXI.

No A 787.—November 1, 1928.

PAGE

Solid Dipleidoscope Prisms.—Supplement. By C. V. Boys, F.R.S 1

The Magnetic Susceptibility of Single Crystals of Zinc and Cadmium. By J. C McLennan, F.R.S., R. Ruedy and E. Cohen 9

On Aerofoils of Small Thickness. By H. Jeffreys, F.R.S 22

The Effect of Electric Fields on the Emission of Electrons from Conductors By A. T. Waterman. Communicated by O. W. Richardson, F.R S... 28

An Ampere Meter for measuring Alternating Currents of very High Frequency By E. B. Moullin. Communicated by R. V. Southwell, F R S. 41

The Analysis and Prediction of Tidal Currents from Observations of Times of Slack Water. By A. T. Doodson. Communicated by J. Proudman, F R.S 72

The Theory of Metallic Corrosion in the Light of Quantitative Measurements.—Part II. By G. D. Bengough, J. M. Stuart and A. R. Lee. Communicated by H. C. H. Carpenter, F.R.S. (Plate 1) . ... , ... 88

The Change in Elastic Properties on Replacing the Potassium Atom of Rochelle Salt by the Ammonium Group. By W. Mandell. Communicated by O. W. Richardson, F.R.S. 122

The Determination of the Piezo-Electric Moduli of Ammonium Seignette Salt. By W. Mandell. Communicated by O. W. Richardson, F.R.S. 130

The Catalytic Decomposition of Gaseous Acetaldehyde at the Surface of Various Metals. By P. C. Allen and C. N. Hinshelwood. Communicated by Sir Harold Hartley, F.R.S. 141

A Quantitative Study of the Reflexion of X-Rays by Sylvine By R W. James and G. W. Brindley. Communicated by W. L. Bragg, F R.S. 155

The Mobility of Ions in Air. Part IV.—Investigations by Two New Methods. By A. M. Tyndall, L. H. Starr and C. F. Powell. Communicated by A. P. Chattock, F.R.S. 172

The Mobility of Ions in Air. Part V.—The Transformation of Positive Ions at Short Ages. By A. M. Tyndall, G. C. Grindley and P. A. Sheppard. Communicated by A. P. Chattock, F.R.S... 185

A Mechanical Method for solving Problems of Flow in Compressible Fluids. By G. I. Taylor, F.R.S., and C. F. Sharman. (Plate 2) 194

PAGE

The Soft X-Ray Levels of Iron, Cobalt, Nickel and Copper. By O. W. Richardson, F.R.S., and F. C. Chalklin 218

An Investigation of Some Banded Structures in Metal Crystals. By C. F. Elam, with an Appendix by G. I. Taylor, F.R.S (Plate 3) 237

The Change in Lattice Spacing at a Crystal Boundary. By J. E. Lennard-Jones and B. M. Dent. Communicated by S. Chapman, F.R.S. 247

The Influence of the Earth's Magnetic Field on Electric Transmission in the Upper Atmosphere. By S Goldstein. Communicated by Sir Joseph Larmor, F.R.S. 260

The Thermal Conductivities of Carbon Monoxide and Nitrous Oxide. By H. Gregory and C T. Archer. Communicated by H. L. Callendar, F.R.S 285

Studies on Fluorescence and Photosensitization in Aqueous Solution. I — Introduction. By W West, R. H. Müller and E. Jette. Communicated by James Kendall, F.R.S. 294

Studies on Fluorescence and Photosensitization in Aqueous Solution II.—Fluorescence in Aqueous Solution By E. Jette and W West Communicated by James Kendall, F.R.S 299

Studies on Fluorescence and Photosensitization in Aqueous Solution. III — Photosensitization and Fluorescence. By R H. Müller. Communicated by James Kendall, F.R.S. 313

The Spectrum of Doubly-Ionised Nitrogen (N III). By L J. Freeman. Communicated by Prof. A. Fowler, F R.S. 318

The Sorption of Carbon Tetrachloride at Low Pressures by Activated Charcoals. Part I.—Apparatus and Method. By R. Chaplin. Communicated by F. G. Donnan, F.R.S. 344

The Structure of Topaz [Al (F, OH)]$_2$ SiO$_4$. By N. A. Alston and J. West. Communicated by W. L. Bragg, F.R.S. 358

The Rate of Emission of Alpha Particles from Radium. By H. J J. Braddick and H. M. Cave Communicated by Sir Ernest Rutherford, P.R.S. 367

The Structure of the Band Spectrum of Helium.—V. By W. E. Curtis and A. Harvey. Communicated by T. H. Havelock, F.R.S. 381

Fluid Motion in a Curved Channel. By W. R. Dean. Communicated by S. Chapman, F.R.S. 402

Some Remarks concerning the Production and Absorption of Soft X-Rays and Secondary Electrons. By E. Rudberg. Communicated by O. W. Richardson, F.R.S. 421

Hyperfine Structure in the Arc Spectrum of Cæsium and Nuclear Rotation. By D. A. Jackson. Communicated by F. A. Lindeman, F.R.S. 432

The Internal Conversion of Gamma-Rays.—Part II. By B. Swirles. Communicated by R. H. Fowler, F.R.S 447

PAGE

The Increase of Thermionic Currents from Tungsten in Strong Electric Fields. By R. S. Bartlett. Communicated by O. W. Richardson, F.R.S. 456

On the Rate at which Particles take up Random Velocities from Encounters according to the Inverse Square Law. By L. H. Thomas. Communicated by R. H. Fowler, F.R.S. 464

The Refractive Index of Quartz. By W. R. C. Coode Adams. Communicated by T. M. Lowry, F.R.S. 476

No. A 788.—December 1, 1928.

The Mode of Formation of Neumann Bands. Part I.—The Mechanism of Twinning in the Body-Centred Cubic Lattice. By S. W. J. Smith, F.R.S., A. A. Dee and J. Young 477

The Mode of Formation of Neumann Bands. Part II.—The Evidence that the Bands are Twins. By S. W. J. Smith, F.R.S., A. A. Dee and J. Young. (Plates 4–7) 486

The Mode of Formation of Neumann Bands. Part III.—The Movement from which the Twinning Results. By S. W. J. Smith, F.R.S., A. A. Dee and J. Young. (Plates 8 and 9) 501

The Wave Pattern of a Doublet in a Stream. By T. H. Havelock, F.R.S. 515

A Symmetrical Treatment of the Wave Equation. By A. S. Eddington, F.R.S. 524

On the Principle of Least Action in Wave Mechanics. By J. M. Whittaker. Communicated by E. T. Whittaker, F.R.S. 543

The Strength of Tubular Struts. By A. Robertson. Communicated by R. V. Southwell, F.R.S. (Plates 10–14) 558

On the Stability of Running of Locomotives. By F. W. Carter. Communicated by A. E. H. Love, F.R.S. 585

The Spacial Distribution of Photoelectrons produced by X-Rays. By E. J. Williams, J. M. Nuttall and H. S. Barlow. Communicated by W. L. Bragg, F.R.S. 611

The Effect of the Image Force on the Emission and Reflexion of Electrons by Metals. By L. W. Nordheim. Communicated by R. H. Fowler, F.R.S. 626

The Thermal and the Electrical Conductivity of a Copper Crystal at Various Temperatures. By W. G. Kannuluik and T. H. Laby. Communicated by Sir Ernest Rutherford, F.R.S. 640

The General Sampling Distribution of the Multiple Correlation Coefficient. By R. A. Fisher. Communicated by Sir James Jeans, Sec. R.S. 654

The Scattering Power of a Bare Nucleus according to Wave Mechanics. By G. Temple. Communicated by S. Chapman, F.R.S. 673

The New Metric of Einstein and the Wave Equation. By H. T. Flint. Communicated by O. W. Richardson, F.R.S. 676

OBITUARY NOTICES.

PAGE
Christian Felix Klein (with portrait) i
Hendrik Antoon Lorentz (with portrait) xx
Theodore William Richards (with portrait) xxix
Sir John Isaac Thornycroft (with portrait) xxxv

Index xxxix

Minutes of Meetings of November 1, 8, 15, 1928.

PROCEEDINGS OF
THE ROYAL SOCIETY.

SECTION A.—MATHEMATICAL AND PHYSICAL SCIENCES.

Solid Dipleidoscope Prisms.—Supplement.

By C. V. BOYS, F.R.S.

(Received June 29, 1928.)

All the prisms described in Sections 1 to 8* are essentially dipleidoscopic about a plane—that is, the two images of the object seen are, one the same as the object by double internal reflection and the other a looking-glass reversed image.

It is possible to make a block of glass of such a shape that two images are seen, each of which is a looking-glass image of the object, but one inverted with respect to the other. Such a glass is dipleidoscopic about an axis. The genesis of such a glass follows naturally from any one of the solid dipleidoscope prisms already described. Taking, for instance, the polarising prism of fig. 4, this outline is shown again in the middle figure of fig. 12, but as the two end views indicate, it is no longer a prism but a tetrahedron, with the angle between the faces AA′B and AA′C, conveniently called the upper angle at A, equal to 67° for ordinary crown glass and its opposite angle along BC equal to 90°, and symmetrically disposed in relation to the edge AA′. Also, of course, the angle between the line BC and the plane AA′C is half the angle at A. If the line bisecting both the angle at A and the transverse angle along BC is directed to an object, then this may be seen by single reflection from AA′B as a looking-glass reversed image, and by treble reflection from the three faces ABC, A′BC and AA′C also as a looking-glass reversed image but inverted; and the two images will overlap so that the point to which the aforesaid bisecting line is directed is the one which is superposed upon itself. Whichever primitive prism may be chosen from which to develop a corresponding tetrahedron, the two reflections from the "roof" faces must be total, and so where the angle at A is less than 74°, already referred to at the end of Section 4, there is no occasion to silver the roof-shaped base. On the other hand, where the angle at A is greater than 74°, the total reflections from the base and long side will cause the trebly reflected light to be far brighter than the externally reflected light, as indicated in the discussion in Section 4 of the properties of the squat prism; but equality can be attained by the method of strip silvering described in the fourth paragraph from this.

---

* Roy. Soc. Proc., A, vol. 119, p. 489 (1928).

As the limit of definition of a tetrahedron dipleidoscopic about an axis should, in general, be equal in all directions, and as the beam reflected from the long face AA′C of a dipleidoscopic prism or tetrahedron is of necessity parallel to the base BC—*i.e.*, to the direction of projection of the right-hand end view in fig. 12, the dotted circle in this view and the dotted circle in the other end view show the greatest beam of circular cross-section that can be transmitted through the tetrahedron, and the dotted tangent lines to this circle and the corresponding dotted lines in the side view indicate the limits of the tetrahedral ears and of the tetrahedral toe, which are redundant. The total distance traversed within the tetrahedron by any ray of light is equal to the length AA′ in the case of the polarising tetrahedron, while in the case of the squat tetrahedron it is equal to BC.

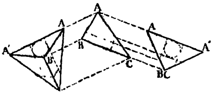


Fig. 12.

As the angle at A of the solid dipleidoscope prism from which the corresponding tetrahedron is developed is made larger, this tetrahedron becomes more symmetrical until, when the angle at A is also a right-angle, the tetrahedron becomes perfectly symmetrical about a single axis which bisects the two opposite dihedral angles of 90°, while the four equatorial angles are 60° each, and the tetrahedron becomes dipleidoscopic in four directions at the same time.

A transit instrument with a tetrahedron suitably mounted in the place of a prism could be used with a telescope as a mere viewing and magnifying machine to determine both time and declination; or, if the tetrahedron were mounted in a cage with a declination axis and circle carried on a polar axis and a viewing telescope were provided, such a tetrahedron would provide the direction factor of an equatorial. I mention these not because I anticipate that either would be a convenient construction, but rather to illustrate the capacity of the dipleidoscopic tetrahedron.

In Sections 3, 4 and 9 I have shown how, by the use of silver or of oil, or simply by the use of plain glass, I have been able to attain sufficient equality in the two images for good observation, and I have shown what fraction of the incident light is available. Where, however, it might be desired to make the equality more perfect, or, as is more likely, to increase the amount of light available, this can be done where the angle A exceeds 74° by the use of strip silvering without diminution of definition. When the angle A exceeds 74° the second internal reflection is total. Unless, therefore, the greater part of the light were allowed to leave the prism at its first reflection from the base, the light emergent at the front face would overpower the light reflected from that face. If, however, the middle half of this face, reaching from A to B, is silvered and polished on its outer surface and the whole of the base BC is silvered and painted for protection, then a telescope taking in the whole of the light from the front face will receive eight or nine times as much light as it would do if there were no silvering, and the proportion of the front face strip of silver may be adjusted to give equality. The beams of light with long narrow cross sections, with their length in the direction in which definition is important, are actually better than the complete circular beam for definition in that one sense only, a fact which I utilised in my apparatus for weighing the earth.*

This method of strip silvering is particularly advantageous in the case of the prism of 90°, 45°, 45° used dipleidoscopically. This I dismissed in the first discussion on solid dipleidoscope prisms as being unsatisfactory on account of disparity in illumination and of waste of glass. The first objection is cured by the strip silvering and the second in considerable degree by omitting the redundant glass. This prism has the great convenience that the double beam of light to be observed is perpendicular to the incident beams, and so leads to the design of the particularly convenient transit instrument described later. Further, should it be desired to observe the transit from both the East and West directions simultaneously—*e.g.*, from one with an eyepiece and from the other photographically—this could be done, but in that case the glass which is redundant for observation from one direction is essential for observation from the other, and the prism must be complete.

Similarly, strip silvering could be used with the prismatic astrolobe described in Section 9. The middle half of the upper face would be silvered and polished on the front face, and mercury would be used in the place of oil.

Since writing the description of my proposed modification of the prismatic astrolobe of Claude and Driancourt, I have received from Messrs. Hilger a beautiful prism of 110°, 35°, 35° with the slant faces 1·9 by 1½ inches, and have mounted it on a 1½-inch telescope as illustrated in fig. 11, and with the necessary adjustments. In the short time available I have not been able to get any observations on a real star, but I put a lighted candle on a table at the top of a flight of stairs, and at the foot of the stairs at a point where the candle was at an elevation of 35°, I set up the instrument with an horizon of olive oil. I then observed the two candle flames almost equal in brightness, one inverted, and I saw the fibres in the red spot at the end of the wick in duplicate, and as the candle burned down the two red spots very slowly approached one another, coincided, and separated again about as slowly as the two images of a star between the pole star and the pole would do if observed at latitude 35°. The use of this observation is to show that as this dull red spot could be so well observed by daylight, there is not likely to be much necessity for strip silvering.

---

* 'Phil. Trans.,' A, vol. 186, p. 1 (1895).

I have not referred to the half silvering such as is used in an interferometer on account of its tender character, and because it has no advantage for these observations over the method of strip silvering. Where I have described front surface silvering, I mean this to include such processes as depend on spluttered platinum, galena or other coatings, which may be more permanent even though not so highly reflective as freshly polished silver.

In fig. 13 I have given an outline of design of a new type of transit instrument

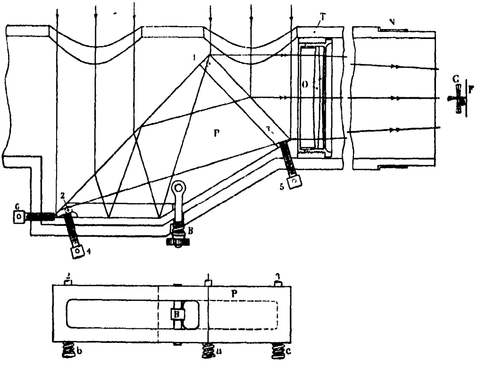


Fig. 13.

which appears to me to be more suitable for time observations of the highest precision than transit instruments of recognised design. The prism P is held in a fixed position in the centre of the length of a steel tube T, resting near its ends in V's. The exact position of the prism is determined by the use of six screws, very much as in the window transit instrument, except that the prism is pressed against the three screws 1, 2, 3, which bear upon one side by three springs, as shown in the plan of the prism below, one opposite to each It is drawn down into contact with the two screws 4, 5, which determine its most important angular position, and against the end screw 6 by the action of a spring-pressed bolt B, as indicated. The use of a hole through a prism to accommodate the transverse bar which is pulled by a bolt is unheard of in optics, but it is convenient in this case, and I believe harmless, as it is outside the region traversed by any rays of light. The two screws 4, 5, which resist this pull bear upon two facets cut in the glass in line with one another, and outside the region traversed by rays of light. The prism is one of 89° 45′ or thereabouts, but the 2 : 1 ratio necessary for a dipleidoscope prism must be carefully maintained. The six screws are held in a massive steel trough welded in a gap in the tube cut to receive it, and two holes are made on the opposite side of the tube to allow the rays from the star to enter. The surfaces of the tube where it rests in the V's are plated with a thick layer of coherent nickel, N, ground as equal and circular on dead centres as modern precision methods allow. Flexure of the tube is avoided by the use of two, three or four bearing rollers (not shown) to carry the greater portion of the weight. Within the tube an object glass, O, of long focus is fixed, and large enough to receive all the rays from the rectangular face of the prism. The prism is so fixed within the steel tube that when the two images of a star coincide the light coming to a focus shall pass axially down the tube or nearly so. The V's are placed east and west and level, exactly as with any transit instrument, and the transit observations of the instrument are used for the purpose of these adjustments exactly as they always are. No piers are required to carry the V's.

The light from the object glass comes to a focus some distance beyond the end of the tube, and it is received upon a cinema film, F, travelling in a direction perpendicular to the movements of the two images. Just in front of the film there is a small piece of plane parallel glass, G, through which the rays pass, which is electromagnetically tilted through a small angle backwards and forwards at each alternate tick of the clock. The space between the end of the steel tube and the photographic receiver is protected from stray light and disturbance by a tubular shield which, except at the time of observation, may be kept rotating if necessary to maintain uniformity of temperature of the air in any cross section. A divided circle is carried by the steel tube with two vernier or pointer rings, each with a spirit level. When an observation is to be made the two pointers are set so that when one or other level shows level the prism is directed to the known altitude of the star. As the star approaches the meridian the two star images on the moving film gradually get nearer together, and ultimately cross, giving rise to two diagonal trails intersecting one another. These cross when the star is one minute or so to the east of the meridian. The steel tube is then lifted by raising and turning gear and reversed in its V's, and a second tubular shield and photographic receiver in the opposite direction is ready, so that when the star is to the same extent past the meridian a crossing of trails on a second film takes place. The two trails are not quite straight diagonal lines, but these contain small steps due to the rocking of the plate by the clock as shown in fig. 14, thus the exact moment of each apparent transit is referred directly to the clock, and the half-way moment between the apparent transits is the observed transit.

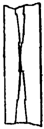


Fig. 14.

It will be seen that with the suggested design the only requirement of absolute character is that the trunnion surfaces should be truly round and equal, and that the V's should be so located that the trunnion axis lies east and west and is level. Next it is essential that the prism should not change its position relative to the trunnion axis between east and west observations. The necessary degree of optical perfection in the prism itself may be difficult to attain, and I discuss this problem later. Given these requirements, it does not matter what the exact angle of the prism is, or whether it is located so that the double beam of light is directed quite axially along the tube. It does not matter if the object glass is exactly central or even if there is any slight motion in its cell between the east and west observations. Further, it does not matter where exactly the rocking plate is situated, or how much it rocks, provided that it is given a minute snappy motion, say, every alternate second, for two seconds is the unit of time of a second's pendulum. Again, it does not matter where exactly the film is placed or the exact direction or speed of its movement, provided that this maintains its uniformity throughout each period of two seconds. For this reason driving clocks with governor control should be avoided, and a simple spring-driven fly be used in preference. Governor control is as pernicious where real uniformity of speed is desired as is ordinary thermostatic control where real uniformity of temperature is required. Such devices give a good average with unknown small spasmodic changes, and should be avoided if real uniformity is essential If the crossing images are not nearly enough axial, the photographic gear can be carried in a ring on a radial carriage, but the discussion of such mere details of design can be deferred to the time when it is settled whether to use hexagon or octagon lock nuts for the six prism screws At the same time it will be convenient to settle the arrangement for ocular observation of the images before they reach the film or while they cross it

Trouble due to flexure with which astronomers are only too familiar is very completely eliminated in the proposed design The prism is placed in the centre of the trunnion axis so that any flexure of this axis should not cause any angular motion of the prism which is the only motion that could matter The extremely rigid steel tube would not bend under its own weight to any appreciable extent The support of nineteen twentieths of its weight by properly positioned rollers will reduce this infinitesimally in proportion and finally if there were any residual effect within the region of observation it would not matter for it would produce equal and opposite effects at the east and west observations Thus while experimentalists are generally pleased when they have managed to arrange for errors to be of the second order of a small quantity conveniently written 0^2 they have not before so far as I know been able to express any of them as 0^4

It is hardly necessary perhaps to point out that the rocking plate overcomes the difficulty or rather futility of marking the film itself by any clock recorder because there would be no means of correlating the clock marks and the unknown position of the star images I mention this because Prof Turner told me that his view was that improvement in transit observations lay in the direction of photographic record I realised that my proposed instrument was peculiarly well adapted for this but I see no satisfactory way of doing this with the usual type of instrument At any rate I have to thank Prof Turner for putting me on to the photographic method

The final question is the production of a prism of the necessary perfection for the material must be optically homogeneous and the faces truly plane and the 2 1 ratio and freedom from pyramidal error should each be within 1′ of perfection

If the prism is made of glass the fact that a low refractive index of 1 51 or so is all that is wanted is all to the good and Sir Charles Parsons considers that this does not present serious difficulty Lord Rayleigh has pointed out to me that crystal quartz is so perfect as to allow definition of the full perfection due to the aperture which at least with the denser glasses is hardly attainable

I had feared the complication of double refraction and of rotation if the refracting edges were not absolutely true to the optical axis, but he has satisfied me by experiment that even with a departure of several degrees there is no rotation of the plane of polarisation, which perhaps I ought to have known. Quartz, strongly twinned and therefore useless for piezo-electric operation, would be as good as pure right- or left-handed quartz if, as I believe is the case, the refractive index of the two kinds is identical, and the axes of the two kinds are parallel. Such quartz may well be available, and it has the further very important advantage of conducting heat in the axial direction twelve times* as well as crown glass, so that optical strain due to temperature change should be far less with quartz than with glass.

Dr. Russell, of Princeton University, with whom I had the advantage of discussing my proposed transit instrument, feared that changing temperature might give rise to serious optical stress, and he thought it might be necessary to enclose that portion of the tube in a thermostatic device. This would entail the addition of two plane windows, which I would rather avoid, and my plan of cutting out the middle half of the prism, except at the ends, as shown in fig. 13, leaving two transparent sides, each of one-quarter of the thickness of the whole prism, I should prefer until experience shows that thermostatic control is imperative. Such a carcase of a prism, whether of quartz or of glass, would pick up changing temperature and suffer in defining power from consequent optical strain to only about one-tenth of the extent of a solid prism in which the unused inner material is retained.

The cooling of the prism could be hurried by exhausting air from the steel tube, except at the time of observation, for the stream of incoming air would bathe the thin walls of the prism.

It is obvious that the use of a film moving several times as fast as the star image will reduce the photographic action on the film, but the discussion of the best proportions is better postponed until I can go into the details of design critically with the astronomers, for which purpose I hope the Royal Astronomical Society will give me an opportunity to meet them.

---

* Kaye and Laby, "Physical and Chemical Constants," p. 55 (1926).

The Magnetic Susceptibility of Single Crystals of Zinc and Cadmium.

By Prof. J. C. McLennan, F.R.S., Dr R. Ruedy, and
Elizabeth Cohen, M.A.*

(Received June 28, 1928.)

During the past five years great progress has been made in the theory of the diamagnetic and paramagnetic properties of matter. Langevin's theory of paramagnetism, long recognised as insufficient, has been replaced by a more general treatment due to B. Cabrera.† The magnetism of metals for which no theory existed at all has been explained by W. Pauli‡ as due to the spin of the free electrons contained in the volume V occupied by the metal. Owing to their enormous concentration in addition to their small mass, the electrons within the metal do not follow the ordinary laws for the gaseous state, in particular, the energy distribution is not Maxwellian. For a Maxwellian distribution, that is, for strongly decreasing energy with decreasing temperature, the electrons could not remain in the free state at low temperature, but would combine with the metallic ions between which they move. The electrons are instead subject to quantisation and the energy is distributed according to the principle of Pauli-Fermi. If ε_k be the energy in the quantum state k, of one of the electrons contained in the volume V, ε_m its spin, then there can be at the most up to $G = 2$ electrons representing this state, and then they must spin in opposite directions. Moreover, if $m\xi$, $m\eta$, $m\zeta$ are the components of the angular momentum of the electron along the three axes of the system of co-ordinates so that $E = \frac{1}{2}mv^2 = \frac{1}{2}m\,(\xi^2 + \eta^2 + \zeta^2)$, then there exist, in the interval between ε and $\varepsilon + d\varepsilon$, $4\pi GV\,(m/h)^3\,v^2dv = 2\pi VGh^{-3}\,(2m)^{3/2}\,\varepsilon^{1/2}\,d\varepsilon$ possible states. Evidently the total energy will be $E = \Sigma n_k\,(\varepsilon_k + \varepsilon_m)$, the only values for n_k being either 0 or 1, with the condition that the total number N of electrons is $N = \Sigma_k n_k$.

Now ε_m, the energy due to the electron spin, is small compared with ε_k and can be neglected in most problems. We may say that the state of the gas is defined if we know the most probable values of n_k for the ε_k, that is, whether n_k is equal to 0 or 1 for a certain ε_k. Owing to the presence of the electronic spin $\varepsilon_s = \pm\, eh/4\pi mc$ (Bohr magneton) the energy distribution will be slightly changed when an external field H is applied. We have for the number of atoms with an orientation parallel to the field

---

* One of the authors, Elizabeth Cohen, was enabled to co-operate in this work by a Studentship awarded to her by the National Research Council of Canada.

† 'J. Physique,' vol. 8, p. 257 (1927).

‡ 'Z. Physik,' vol. 41, p. 81 (1927).

$$\mathrm{N}_{+\epsilon m} = \frac{\mathrm{V}2\pi\,(2m)^{3/2}}{h^3}\int_0^\infty \varepsilon^{1/2}\left(\frac{1}{\mathrm{A}}e^{\frac{\varepsilon+\epsilon m}{k\mathrm{T}}}+1\right)^{-1}d\varepsilon,$$

and for those with anti-parallel orientation

$$\mathrm{N}_{-\epsilon m} = \frac{\mathrm{V}2\pi\,(2m)^{3/2}}{h^3}\int_0^\infty \varepsilon^{1/2}\left(\frac{1}{\mathrm{A}}e^{\frac{\varepsilon-\epsilon m}{k\mathrm{T}}}+1\right)^{-1}d\varepsilon.$$

A magnetic moment results and the magnetic susceptibility is $\chi = \frac{n}{2}\frac{\mu^2}{k\mathrm{T}\log\mathrm{A}}$. As a first approximation

$$\log\mathrm{A} = \frac{h^2}{2mk\mathrm{T}}\cdot\left(\frac{3n}{4\pi\mathrm{G}}\right)^{2/3},$$

hence

$$\chi_0 = \mathrm{G}^{2/3}\left(\frac{3n}{4\pi}\right)^{-2/3}\frac{nm\mu^2}{h^2} = 4\left(\frac{\pi}{6}\right)^{2/3}\mathrm{G}^{2/3}\frac{\mu^2 n^{1/3}m}{h^2},$$

the formula deduced by Pauli for the magnetic susceptibility of the metals.

In order to recognise the influence of temperature upon the susceptibility, we have to take the next closer approximation for the value of log A, namely,

$$\log\mathrm{A} = h^2\left(\frac{3n}{4\pi\mathrm{G}}\right)^{2/3}\left\{1-\left(\frac{2\pi mk\mathrm{T}}{12h^4}\right)^2\left(\frac{3n}{4\pi\mathrm{G}}\right)^{-4/3}\right\}$$

or

$$\frac{1}{\log\mathrm{A}} = \frac{1}{h^2}\left(\frac{3n}{4\pi\mathrm{G}}\right)^{-2/3}\left\{1+\left(\frac{2\pi mk\mathrm{T}}{12h^4}\right)^2\left(\frac{3n}{4\pi\mathrm{G}}\right)^{-4/3}\right\},$$

hence

$$\chi = \frac{n\mu^2 m}{h^2}\left(\frac{3n}{4\pi\mathrm{G}}\right)^{-2/3} + \left(\frac{2\pi mk\mathrm{T}}{12h^6}\right)^2\cdot n\mu^2 m\left(\frac{3n}{4\pi\mathrm{G}}\right)^{-2}$$

$$= \chi_0 + \frac{(4\pi^2\mathrm{G}k)^2}{3^3h^6}\cdot\frac{m^3\mu^2}{n}\cdot\mathrm{T}^2 = \chi_0 + \frac{\mathrm{T}^2}{n}\cdot 10^7.$$

As n is of the order of 10^{22} (atoms per cm.^{3}), it becomes evident that the influence of the temperature is negligibly small. There would be indeed no other property of the metal which is so independent of temperature as its magnetic susceptibility. In view of this result it would seem advisable to remeasure some of the metals for which the older work* has given a change in susceptibility with temperature (Al, Ti, Mn, etc.).

There is, however, occasion for experimental work upon a more important point which Pauli and later Sommerfeld† had to neglect when treating the free electrons inside metals; the presence in the metal of the remainder of the atoms. If the electrons would be perfectly free particles, the distribution of the metallic ions would be indifferent. It is well known, however, that metallic and thermic conductivity differ in single metallic crystals according to the direction considered, unless they belong to the cubic system. Formally we may try to account for this by assuming that the velocity distribution around a point inside the metal is not of spherical symmetry.

---

* Stoner, 'Magnetism and Atomic Structure' (1926).

† 'Z. Physik,' vol. 47, p. 1 (1928).

In this connection it is equally important to study the magnetic properties of single metallic crystals. In the new theory the magnetic properties are of a more simple nature than the electrical or thermal characteristics, being entirely a consequence of the spin of the electron and the Pauli-Fermi distribution of energy, two principles which have found wide application in atomic physics. The present paper gives the results of a study of the single crystals of Zn, Cd, and Hg. The susceptibility of the metals of the fifth column, which may also be obtained in single crystals, is larger, it is true, but they seem to present additional complications not due to atomic susceptibility alone.

The metals Zn and Cd both crystallise in the hexagonal system. The atoms are more widely spaced along the hexagonal axis than in the direction normal to it, the ratio c/a being 1·86 for Zn and 1·89 for Cd. The comparatively large difference between the crystal axes renders these metal crystals interesting, and Zn and Cd have received a great deal of attention in the past few years by workers in various fields. Some of the more important constants parallel or perpendicular to the hexagonal axis, due to Bridgman and Meissner, Grueneisen and Goens,* are summarised in Table I.

Table I.

Initial linear compressibility at 30° C	Specific resistance at 20° C.	Linear expansion at 20° C.
Zn ∥ 12·98 . 10^{-7}	6·13 . 10^{-6} B 5·83 M	57·4 . 10^{-6}
⊥ 1·95	5·91 B 5·39 M	12·6
Cd ∥ 18·3	8 3 B 6·8 M	52·5 (G)
⊥ 2·1	6·8 B 6·5 M	20 2

---

* Bridgman, 'Proc. Am. Acad.,' A, vol. 60, p. 305 (1925); Meissner, 'Z. Physik,' vol. 38, p. 647 (1926); Gruneisen and Goens, 'Z. Physik,' vol. 26, pp. 235, 250 (1924).

In addition a study of the photoelectric effect* has led to the conclusion that the emission of electrons varies with the orientation and is largest for the surface parallel to the basal plane. This is at the same time the plane of easiest cleavage of the crystal and can usually be recognised by its high lustre and fine triple striations.

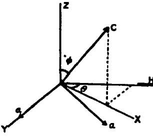


FIG. 1.

If now a, b, c be the principal susceptibilities of a crystal, then the susceptibility χ_r corresponding to a definite direction, of which the direction cosines are α, β, γ, is given by†

$$a\alpha^2 + b\beta^2 + c\gamma^2 = 1/r^2 = \chi_r.$$

As the crystals considered here belong to the hexagonal system, we must expect a and b taken at right angles to the hexagonal axis to be equal, so that

$$/r \quad (\alpha^2 + \beta^2)\,a + \gamma^2 c.$$

Consider the values of χ_r in a plane perpendicular to the cylindrical axis. Let ϕ be the angle between the axis of the cylinder and the principal axis of the crystal, θ the angle between the plane containing z and c (fig. 1) and the direction of χ_r, then:

$$\chi_r = a\,(\sin^2\theta + \cos^2\phi\cos^2\theta) + c\sin^2\phi\cos^2\theta.$$

If χ_0 is the susceptibility for $\theta = 0$, χ_{90} the value for $\theta = 90°$ is

$$a = \chi_{90}, \quad c = (\chi_0 - \chi_{90}\cos^2\phi)/\sin^2\phi.$$

The variations of χ_r with θ will therefore be

$$\chi_r = \chi_0 + (\chi_{90} - \chi_0)\sin^2\theta.$$

Preparation of Crystals.

The crystals used in this work were made by the method of very slowly lowering the molten metal,‡ contained in a mould of pyrex out of an electric furnace. If the rate of lowering be at least as slow as the velocity of crystallisation, the solidification starts at the bottom of the tube and proceeds along its axis. The glass moulds were made of a special shape designed to increase the probability of obtaining a single crystal through the whole length of the tube (fig. 2). The end A was drawn into a very narrow point so that when the melt begins to solidify, there would be room for only one crystal to form. The orientation of this crystal may not be the most favourable for growth, and in that case it

---

* Linder, 'Phys. Rev.,' vol. 30, p. 649 (1927).

† Jackson, 'Phil. Trans.,' vol. 226, p. 126 (1926).

‡ Bridgman, *loc. cit.*

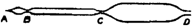


FIG. 2.

is likely that other crystals may form farther along the tube. A fine constriction at B acts as a filter letting through the one germ whose orientation is most favourable for growth.

The purity of the metal was of great importance, both from the point of view of preparing the crystal, dirt particles, such as specks of oxide, forming nuclei around which new crystals might start, and in determining the magnetic susceptibility of the crystals, very small traces of iron causing considerable effect on the result. Chemically pure metals were used and as an extra precaution these were slowly distilled in a good vacuum. The tubes were cleaned by washing out with dilute acid and distilled water. They were then dried and oiled by filling with Nujol, which was washed out by several fillings of ether. The tube was then inverted and the oil driven out with a flame. This process leaves a trace of grease on the walls of the glass and so makes it possible to remove the glass later.* If both the metal and the glass were perfectly clean they might stick to each other.

To prepare a crystal, the metal was placed in CD and the end D of the tube was then drawn out, joined to a stop-cock and thence to a Hyvac pump. The tube was then inserted up to C in the furnace in a horizontal position and the furnace allowed to come to a temperature well above the melting point of the metal. The metal was now melted with a bunsen burner, and the molten metal was rocked back and forth in CD for 10 to 15 minutes to get rid of any occluded gases in the metal. Tube and furnace were then rotated to a vertical position and the metal filtered through C to fill the tube. The tube was allowed to cool slowly and was then sealed off at C or D. It was now attached to the lowering device and inserted in a second furnace so that the bottom of the tube, A, was at about the middle of the furnace. The centre of this furnace was kept well above the melting point of the metal under consideration.

The lowering device adopted for making these crystals consisted of an ordinary alarm clock, the setting screw of which had been replaced by a drum of the required diameter, the cirumference of the drum giving the lowering per hour; 1·6 cm. for cadmium and 2·9 cm. for zinc. This drum was wound with heavy linen thread which was attached to the tube by a fine copper wire.

---

* Bridgman, *loc. cit.*

The construction of the furnaces used is indicated in fig. 3, which shows the arrangement for lowering the crystal. A is a quartz tube 2½ cm. in diameter and 30 cm. in length (for one of the furnaces an alundum tube was used), which was closely wound with nichrome wire, B. The wire was held in place by alundum cement, and thermally insulated by asbestos packing C. One furnace was fastened rigidly in a vertical position; a second furnace was pivoted at the position PP, so that it could be rotated into a horizontal position when desired.

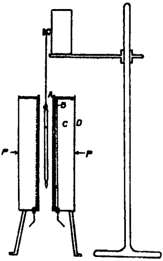


FIG 3

After lowering the tubes the metal was removed from the glass by scratching the tube with a diamond glass cutter and then, in the case of zinc, gently pressing between the jaws of a vice For cadmium the scratched glass was cracked off, holding it in a fine oxy-gas flame and then immersing it in water.

The piece of metal was examined in a beam of light. If the casting were a single crystal the whole surface would, when it was rotated, flash up at one position, as it would be made up of microscopic pits, the sides of which were planes of the natural crystals comprising the casting.

As zinc breaks quite readily along its basal plane it was easy to determine the position of the axis of the crystal relative to the axis of the cylinder. To prevent deformation, the crystal, which is quite soft, was dipped into liquid air before being broken. An X-ray photograph taken* of a thin layer chipped off the zinc crystal showed definitely that the cleavage plane was the basal plane of the crystal (fig. 4). As cadmium crystals did not break without the metal being distorted, the orientation was determined by following the method described by Bridgman. The crystal, threaded through a sphere, was rotated in a beam of light until the light was reflected. A mirror was then held in such a position as also to reflect the light. The back of the mirror, wet with printers' ink, was then touched to the sphere. In this way the positions of the crystal faces could easily be determined.

---

* Through the kindness of Mr. W. G. Plummer, M.Sc.

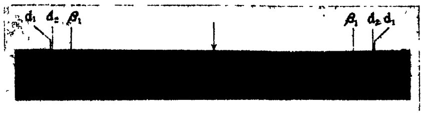


Fig. 4.—Single Crystal of Zinc. Cu_k radiation, reflected from (0001) face. Distance from crystal to film, $5{\cdot}73_5$ cm.

Line.	Distance from line to centre on film (cm.)	Spacing in Å of (0002) plane.
$Cu_{k_{\alpha_2}}$	$4{\cdot}220$	$2{\cdot}47_1$
$Cu_{k_{\alpha_1}}$	$4{\cdot}202$	$2{\cdot}47_4$
$Cu_{k_{\beta_1}}$	$3{\cdot}671_5$	$2{\cdot}47_3$

Theoretical value (0002) spacing = 2·471 Å.

Measurement of Susceptibility.

(The crystals prepared were of about 0·5 cm in diameter and about 6 cm. long.)

The magnetic susceptibility of each metal was determined by the Gouy method (see Stoner, 'Magnetism and Atomic Structure'), the crystal being weighed in the presence and absence of a strong magnetic field The specimen was suspended from one arm of a balance so that its lower end hung between the pole pieces of an electromagnet in the position of uniform field. As the castings were all sufficiently long to justify neglecting the effect of the field at the other end, the volume susceptibility is given by $k = 2g/A \cdot P/H^2$, where A is the area of cross section of the specimen and P is the pull in grams caused by a field H between the pole pieces. In each case a small correction had to be made for the effect of the brass carrier. This was done by taking a set of readings after removing the specimen from the carrier. The carrier was constructed so that the crystal could be rotated about a vertical axis, the position being read on a horizontal circular scale. Readings were taken every 15° for a definite current in the electromagnet, and the results plotted on graph paper. The positions of maximum and minimum effect were noted and readings taken at these positions for currents in the electromagnet varying from 10 to 22 amps., *i.e.*, 8500 to 11,600 gauss approximately.

The field strength was determined by weighing a single loop of wire bearing a current in the magnetic field. The pull caused per ampere by the magnetic field on the wire was squared and plotted against the current in the electromagnet so that the value of this fraction, which is proportional to the square of the magnetic field, could be read off for any current.

Now $\mathrm{H} = \frac{10g}{l}\frac{p_i}{i} = \frac{10g}{l}\mathrm{I}$, where p_i is pull due to magnetic field H on a wire of length l cm. carrying a current i amps. perpendicular to H and p_i.

Therefore the volume susceptibility

$$k = \frac{2g}{\mathrm{A}} \cdot \frac{\mathrm{P}}{100g^2/c^2} \cdot \mathrm{I}^2 = \frac{l^2}{50g\mathrm{A}} \cdot \frac{\mathrm{P}}{\mathrm{I}^2},$$

or the specific susceptibility $\chi = k/d = f\mathrm{P}/\mathrm{I}^2$, where $f = l^2/50\mathrm{A}d$ and d = density of metal.

In calculating the specific susceptibility we used Owen's* method to correct for the iron impurity. Owen gives the relation between field strength and susceptibility to be

$$\chi_{\mathrm{H}} = \chi_{\infty} + \sigma/\mathrm{H},$$

where χ_{H} is the susceptibility at a field H, χ_{∞} and σ are constants, χ_{∞} being the susceptibility at $\mathrm{H} = \infty$, *i.e.*, when the iron effect tends to zero, and σ depends on the iron content. This may be written in the form $\mathrm{P} = \mathrm{TI}^2 + \mathrm{FI}$, where $\mathrm{T} = \chi/f$, $\mathrm{F} = l\sigma/10gf$. In these experiments the variation in the field strength was not very large—8500 to 11,600 gauss approximately—and the curves obtained by plotting P against I^2 were found to be straight lines. These might be considered as tangents to the curve at the mean value of I^2 used. The tangent is given by $\mathrm{P} = [\mathrm{T} + \frac{1}{2}\mathrm{F}/\mathrm{I}']\,\mathrm{I}^2 + \frac{1}{2}\mathrm{FI}'$, so that if the equation of the curve is given by $\mathrm{P} = \mathrm{RI}^2 + \mathrm{S}$, then $\mathrm{T} = \mathrm{S} - \frac{1}{2}\mathrm{F}/\mathrm{I}'$, $\mathrm{F} = 2\mathrm{S}/\mathrm{I}'$, so that $\mathrm{T} = \mathrm{R} - \mathrm{S}/\mathrm{I}'^2$.

Results.

The following tables give the results obtained for one of the zinc crystals measured, illustrating the way in which the susceptibilities were obtained in each case.

---

* 'Ann. Physik,' vol. 37, p. 657 (1912)

Table II.

Readings on horizontal circular scale.	Pull due to 15·5 amps. in electromagnet × (−10⁴).	Readings on horizontal circular scale.	Pull due to 15·5 amps. in electromagnet × (− 10⁴).	Mean pull positions rotated through 180° × (− 10⁴).
°	gms.	°	gms.	gms.
0	165	180	161	163
15	160·5	195	156	158
30	153·5	210	152	153
45	153	225	151	152
60	159	240	161·5	160
75	172	255	169	170
90	181·5	270	179	180
105	188·5	285	191	190
120	200	300	197	198
135	198	315	198	198
150	188	330	191·5	190
165	177·5	345	181·5	179

From the curve (fig. 5) we see that the maximum susceptibility is given for scale reading 127° and the minimum for 37°.

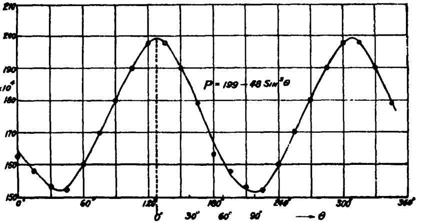


Fig. 5.

Table III.—For Position of Minimum Susceptibility 37°.

Current in electromagnet E.	Pull due to E $\times (-10^4)$.	Pull corrected for effect of carrier P $\times (-10^4)$.	I^2.
amps.	gms.	gms.	
9·0	99	104	2 50
10·3	111	117	2 805
12·0	127	134	3·185
14·2	143	152	3 605
16 4	158	168	3·975
18 1	168	179	4·21
20 0	177	189	4·45
9 8	108	113	2·695
11 1	121	128	2·99
13 5	138	146	3·48
15·3	152	162	3·80
17 8	168	178	4·11
19 0	173	184	4·33
21 0	183	196	4 57

Table IV.—For Position of Maximum Susceptibility 127°.

Current in electromagnet E	Pull due to E $\times (-10^4)$.	Pull corrected for effect of carrier P $\times (-10^4)$.	I^2
amps.	gms.	gms	
9 1	133	138	2 53
10 2	148	154	2·83
12 2	170	177	3·32
14 2	188	196	3·61
16·2	206	216	3·945
18·2	220	231	4 225
20 1	233	245	4·46
9 1	129	134	2·52
11 0	152	158	2·965
13 2	157	185	3·425
15·3	197	206	3 80
17 3	211	222	4·105
19 6	230	242	4 40
21·0	240	252	4 57

P was then plotted against I^2 and the best straight line drawn through the points (fig. 6). In drawing this line consideration was taken of the fact that the constant of the equation, which depends on the iron content of the crystal, should probably be the same for the two positions.

The equations of the curves are

$$(1) \quad P = -445 \, . \, 10^{-5} \, I^2 + 75 \, . \, 10^{-5}$$

and

$$(2) \quad P = -562 \, . \, 10^{-5} \, I^2 + 75 \, . \, 10^{-5}.$$

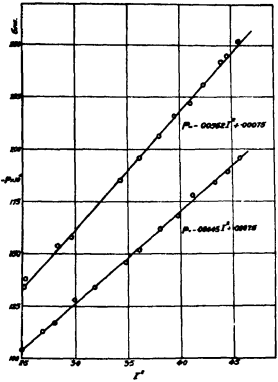


FIG. 6

Now $T = R - S/I'^2$,

so for (1) $T = -445 \,.\, 10^{-5} - 75 \,.\, 10^{-5}/3 \cdot 5 = -446 \,.\, 10^{-5}$.

Now $A = \pi \,.\, (0 \cdot 307)^2$ cm.^{2}, $l = 1 \cdot 79$ cm., $d = 7 \cdot 17$.

Therefore

$$\chi_\infty = 30 \cdot 5 \,.\, 10^{-6}\, T = -0 \cdot 142 \,.\, 10^{-6}.$$

Similarly for (2)

$$T = -583 \,.\, 10^{-5} \text{ and } \chi_\infty = -0 \cdot 178 \,.\, 10^{-6}.$$

$F\, I' = l\sigma/10gf \,.\, \tfrac{1}{2} I'$ we get

$$\sigma = 14 \,.\, 10^{-5}$$

or the amount of free iron per gramme of zinc

$$= 0 \cdot 6 \,.\, 10^{-6} \text{ gm.}$$

In this way the magnetic susceptibility of single crystals of zinc and cadmium were measured and the results obtained are given in the following table :—

Table V.

Crystal.	ϕ.	χ_ϕ $\times (-10^6)$.	$\chi_{90} = a$ $\times (-10^6)$.	$c = (\chi_\phi - \chi_{90} \cos^2\phi)/\sin^2\phi$ $\times (-10^6)$.
	°			
Zinc 1	90	0 184	0·139	0·184
2	90	194	151	194
3	65	177	141	185
4	62	179	143	189
5	90	196	149	196
		Mean	145	190
Cadmium 1	71	0 254	0 164	·0265
2	90	259	156	259
3	90	265	160	265
4	90	254	162	254
		Mean	160	261

From these values we would expect the magnetic susceptibility of an isotropic aggregate of zinc crystals arranged at random to be $\frac{1}{3}(2a + c) = -0{\cdot}160 \,.\, 10^{-6}$. Owen (*loc. cit.*) gives $-0{\cdot}155 \,.\, 10^{-6}$ as the best value. For cadmium the above results give an average susceptibility of $-0{\cdot}194 \,.\, 10^{-6}$. Owen gives $-0{\cdot}18 \,.\, 10^{-6}$.

Mercury.

An attempt was made to grow single crystals of mercury for the purpose of studying its magnetic properties. Work done on the crystal structure of mercury shows two different forms. McKeehan and Cioffi,* working at a temperature of − 115° C., found that mercury crystallised as a rhombohedral crystal. Alsen and Aminoff,* using solid carbon dioxide, obtained the solid mercury in the form of hexagonal crystals. If a casting of mercury were a single crystal of the hexagonal type with its principal plane parallel or inclined at a small angle to the axis of the cylinder, we would expect, in analogy with Cd and Zn, a considerable change in the magnetic susceptibility as the crystal is rotated in the magnetic field. If the crystal were rhombohedral, little change would be expected.

The principle applied in attempting to make mercury crystals was the same as that used for the zinc and cadmium crystals. In this case the metal was lowered from room temperature into a tube surrounded by liquid air. The apparatus is shown in fig. 7. It consists essentially of a brass tube A which is cut away at PP to fit over the pole pieces of the electromagnet. The lower section of the tube, which was surrounded by liquid air, was sealed into a second brass tube, the small air space B serving to prevent the inside of the tube from quite reaching liquid air temperatures. Iron cylinders, CC, are held in position 3½ cm. from the pole pieces by the brass tubes DD. The iron cylinders ensure the value of the field being zero at the ends of the glass tube; air at very low temperatures is very paramagnetic and even the relatively small field at this distance causes an appreciable effect. A tube of non-conducting material G prevents frost from forming across the mouth of the tube. The tube was attached to the clockwork as shown in the diagram and allowed to lower for about 10 hours (rate of lowering, 1·6 cm. per hour). During this time the level of the liquid air in F was maintained constant. The tube was then removed from the clock and attached to the pan of the balance. Weighings were then taken for different readings on the horizontal scale for a current of 15·5 amps. in the electromagnet.

---

* See Wyckoff, 'Structure of Crystals,' p. 242 (1924).

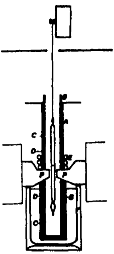


Fig. 7.

As the solid mercury prepared in this way in four different experiments was rotated about its vertical axis the pull caused by the magnetic field remained constant. This result might be taken to indicate that the metal was not in the form of a single crystal. Or, as the method of preparing the mercury crystals used was similar to the one for obtaining Cd or Zn crystals, it might be taken to indicate that the mercury was crystallised and in a rhombohedral form. When we examined the tube containing the solid mercury in a beam of light, indications of a reflection position were noted, but as the glass frosted over very rapidly it was difficult to be certain that the metal was in the form of a single crystal.

Summary.

The magnetic constants of single crystals of Zn and Cd have been determined. For the magnetic susceptibility $\chi_{\parallel}$ (parallel) and $\chi_{\perp}$ (normal) to the hexagonal axis, the results are

$$\text{Cd} \ldots\ldots \chi_{\parallel} = 190 \,.\, 10^{-6}, \qquad \chi_{\perp} = 145 \,.\, 10^{-6}.$$
$$\text{Zn} \ldots\ldots \chi_{\parallel} = 261 \,.\, 10^{-6}, \qquad \chi_{\perp} = 160 \,.\, 10^{-6}.$$

With mercury the results obtained lent support to the view that this metal crystallised in a rhombohedral form.

On Aerofoils of Small Thickness.

By Harold Jeffreys, M.A., D.Sc., F.R.S., St. John's College, Cambridge.

(Received August 10, 1928.)

Theories giving the properties of the general aerofoil of small camber and negligible thickness have been given by Munk,[*] Birnbaum, and Glauert.[†] Allowance for some thickness is desirable, however, since for most aerofoils the thickness is comparable with the camber. The method used is an extension of that of Munk and of Glauert in his first account.

1. If the circle C given by $|z| = a$ is transformed by the equation

$$z' = z + a^2/z, \tag{1}$$

we get the two sides of a flat plate C′ extending from $2a$ to $-2a$ and back. If we first transform the circle by

$$\zeta = z\{1 + \Sigma\, \mathrm{A}_n\, (a/z)^n\}, \tag{2}$$

where the A_n's are small complex quantities, we get a curve S differing slightly from a circle, and if we then put

$$\zeta' = \zeta + a^2/\zeta, \tag{3}$$

we get a curve S′ in the ζ' plane differing slightly from a flat plate. If we write on S

$$\zeta = a\,(1 + r)e^{i\theta}, \tag{4}$$

where r is real and small,

$$\zeta' = 2a\,(\cos\theta + ir\sin\theta) \tag{5}$$

to the first order in r; whence

$$\xi' = 2a\cos\theta\,;\ \eta' = 2ar\sin\theta. \tag{6}$$

Also, if on C

$$z = ae^{i\phi}\,; \tag{7}$$

$$x' = 2a\cos\phi\,;\quad y' = 0. \tag{8}$$

Comparing (4) and (7) with (2), we have

$$(1 + r)\,e^{i(\theta-\phi)} = 1 + \Sigma\,\mathrm{A}_n e^{-ni\phi}, \tag{9}$$

whence r and $\theta - \phi$ are small; and if we put

$$\mathrm{A}_n = \mathrm{B}_n + i\mathrm{C}_n \tag{10}$$

---

[*] 'Nat. Adv. Ctee. for Aeronautics, Washington,' Rept. 142 (1922).

[†] 'R. and M.,' No. 910 (1924); 'Aerofoil and Airscrew Theory' (1926).

we get, to the first order,

$$r = \Sigma\,(B_n \cos n\phi + C_n \sin n\phi), \tag{11}$$

$$\theta - \phi = \Sigma\,(C_n \cos n\phi - B_n \sin n\phi). \tag{12}$$

Hence by Fourier's theorem

$$B_n = \frac{1}{\pi}\int_0^{2\pi} r \cos n\phi \, d\phi = \frac{1}{\pi}\int_0^{2\pi} \frac{\eta'}{2a}\,\frac{\cos n\phi}{\sin \phi}\, d\phi, \tag{13}$$

$$C_n = \frac{1}{\pi}\int_0^{2\pi} r \sin n\phi \, d\phi = \frac{1}{\pi}\int_0^{2\pi} \frac{\eta'}{2a}\,\frac{\sin n\phi}{\sin \phi}\, d\phi, \tag{14}$$

except

$$B_0 = \frac{1}{2\pi}\int_0^{2\pi} r \, d\phi = \frac{1}{2\pi}\int_0^{2\pi} \frac{\eta'}{2a}\,\frac{d\phi}{\sin \phi}. \tag{15}$$

It appears that C_n depends on the sums, B_n on the differences, of the values of η' corresponding to values of ϕ adding up to 2π; thus the C_n's depend on the camber of the aerofoil, the B_n's on its thickness. The transformation (2) is now determinate.

When r is represented as a Fourier series, $\theta - \phi$ is given by the conjugate series, which has a considerable literature in the theory of functions of a real variable. It and the corresponding integral have simple applications in electrostatics.

We notice that near $\theta - \phi$ (the trailing edge) we have approximately

$$\frac{\eta'}{2a - \xi'} = \frac{2r}{\theta}, \tag{16}$$

and near $\theta = \pi$ we have

$$\frac{\eta'}{\xi' + 2a} = \frac{2r}{\psi}, \tag{17}$$

where $\theta = \pi - \psi$. Thus r vanishes at $\theta = 0$ because the slope of the section is finite there; but at $\theta = \pi$ (the leading edge) the section and the axis meet at right angles. If the curvature is ρ we have nearly

$$2\rho = \frac{\eta'^2}{\xi' + 2a} = 2ar^2(1 - \cos\theta) = 4ar^2, \tag{18}$$

so that r is finite and equal to $(\rho/2a)^{\frac{1}{2}}$ at the leading edge. This result, of course, depends on the finite curvature there, and therefore is characteristic of the aerofoil with some thickness.

2. The complex potential when the flow at a great distance has velocity U at an angle α to the x axis, and the obstructing body is the circle C, is

$$w = \Phi + \iota\Psi = Ue^{-\iota\alpha} z + Ue^{\iota\alpha}\frac{a^2}{z} + \frac{\iota\kappa}{2\pi}\log\frac{z}{a}, \tag{19}$$

where κ is the circulation. The velocity is to be finite at the trailing edge, where ξ' is a maximum and therefore θ zero. But the velocity vector

$$u - \iota v = \frac{dw}{d\zeta'} = \frac{dw}{dz} \Big/ \frac{d\zeta}{dz}\frac{d\zeta'}{d\zeta}, \tag{20}$$

and the last factor is zero when $\zeta = a$. Hence dw/dz is zero and

$$\mathrm{U}\,e^{-\iota\alpha} - \mathrm{U}e^{\iota\alpha}\frac{a^2}{z^2} + \frac{\iota\kappa}{2\pi z} = 0. \tag{21}$$

But

$$\frac{a}{z} = e^{-\iota\phi} = e^{-\iota\theta}\,e^{\iota(\theta-\phi)},$$

which becomes when $\theta = 0$

$$\exp \iota\Sigma\,\mathrm{C}_n = \exp \iota\beta, \tag{22}$$

say. Hence, from (21),

$$\kappa = 4\pi a\mathrm{U}\sin(\alpha + \beta). \tag{23}$$

By Joukowsky's theorem the lift L of an aerofoil is $\rho\mathrm{V}\kappa$, where V is the velocity at a great distance and ρ the density of the air. When z is great, $d\zeta/dz$, and hence $d\zeta'/dz$, are equal to $1 + \mathrm{B}_0$. If we keep the form (19) the circulation is the same about the circle and its transformed figures, but the change in the scale alters the velocity at a great distance; in fact,

$$u - \iota v = \mathrm{U}e^{-\iota\alpha}/(1 + \mathrm{B}_0) \tag{24}$$

when negative powers of z are neglected; and hence the direction of the flow is the same, but its magnitude is altered, giving

$$\mathrm{U} = \mathrm{V}(1 + \mathrm{B}_0). \tag{25}$$

Applying Joukowsky's theorem we have now

$$\mathrm{L} = 4\pi\rho a\mathrm{V}^2(1 + \mathrm{B}_0)\sin(\alpha + \beta). \tag{26}$$

The lift coefficient is

$$k_{\mathrm{L}} = \frac{\mathrm{L}}{4\rho a\mathrm{V}^2} = \pi(1 + \mathrm{B}_0)\sin(\alpha + \beta). \tag{27}$$

The thickness therefore affects the lift through B_0, which is a positive constant for a given wing.

3. The moment* about the centre of the aerofoil (that is, the point midway between the leading and trailing edges) is M, equal to $\pi\rho$ times the coefficient of ζ'^{-2} in the expansion of $(dw/d\zeta')^2$. Using (20) we have

$$\begin{aligned}\frac{dw}{d\zeta'} &= \frac{\mathrm{U}\,e^{-\iota\alpha} - \mathrm{U}\,e^{\iota\alpha}a^2/z^2 + \iota\kappa/2\pi z}{\{1 - \Sigma(n-1)\,\mathrm{A}_n\,(a/z)^n\}\,(1 - a^2/\zeta^2)} \\ &= \frac{\mathrm{U}e^{-\iota\alpha} - \mathrm{U}e^{\iota\alpha}a^2(1+\mathrm{B}_0)^2/\zeta'^2 + \iota\kappa(1+\mathrm{B}_0+\mathrm{A}_1 a/\zeta')/2\pi\zeta'}{\{1 + \mathrm{B}_0 - \mathrm{A}_2(a/\zeta')^2\}\,(1 - a^2/\zeta'^2)} + O(\zeta'^{-3}),\end{aligned} \tag{28}$$

---

* *Cf.* Glauert, 'Aerofoil and Airscrew Theory,' pp. 81–82.

and the required coefficient is

$$2V^2a^2\left\{(1+A_2)\,e^{-2\iota\alpha}-(1+B_0)^2+\frac{\iota\kappa A_1 e^{-\iota\alpha}}{2\pi a V}\right\}-\frac{\kappa^2}{4\pi^2}(1+B_0)^2. \tag{29}$$

The terms in $1 + B_0$ and κ^2 are real. Hence

$$\begin{aligned}M &= 2\pi\rho V^2a^2 I\left\{(1+B_2+\iota C_2)\,e^{-2\iota\alpha}+2\iota\,(B_1+\iota C_1)\sin(\alpha+\beta)\,e^{-\iota\alpha}\right\}\\ &= 2\pi\rho V^2a^2\{C_2\cos 2\alpha-(1+B_2)\sin 2\alpha+2B_1\sin(\alpha+\beta)\cos\alpha\\ &\qquad +2C_1\sin(\alpha+\beta)\sin\alpha\}\end{aligned} \tag{30}$$

to the first order in the B's and C's. The moment about the leading edge is

$$M+2aL=2\pi\rho V^2a^2\{2\alpha\,(1+2B_0+B_1-B_2)+C_2+2\beta\,(2+2B_0+B_1)\}, \tag{31}$$

where terms of orders α^2, αB_n, αC_n have been retained, but α^3, $\alpha^2 B_n$, and $\alpha^2 C_n$ dropped. If, further, we drop squares and products of the B's and C's, the moment coefficient k_M is

$$k_M=\frac{M+2aL}{(4a)^2\rho V^2}=\tfrac{1}{4}\pi\alpha\,(1+2B_0+B_1-B_2)+\tfrac{1}{8}\pi C_2+\tfrac{1}{2}\pi\beta \tag{32}$$

$$=\tfrac{1}{4}k_L\,(1+B_0+B_1-B_2)+\tfrac{1}{8}\pi C_2+\tfrac{1}{4}\pi\beta. \tag{33}$$

It follows that the moment coefficient at zero lift is independent of the thickness. This was to be expected, for we have neglected the square of the thickness and its product by the camber, so that only the first powers of the B's could enter the formula in any case. If, then, there was a moment coefficient at zero lift it would be the same, to our order, as that for a symmetrical aerofoil of the same thickness, and this is obviously zero.

The thickness affects the slope of k_M against k_L, so that the centre of pressure is displaced along the aerofoil.

The properties of an aerofoil of small thickness are therefore summed up in the numbers β, B_0, $B_1 + B_1 - B_2$, C_2. We have

$$\begin{aligned}\Sigma A_n &= \frac{1}{2\pi}\int_0^{2\pi} r\,d\phi+\Sigma\frac{1}{\pi}\int_0^{2\pi} r\,e^{\iota n\phi}\,d\phi\\ &= \lim_{n\to\infty}\frac{1}{\pi}\int r\,(\tfrac{1}{2}+e^{\iota\phi}+\dots+e^{n\iota\phi})\,d\phi\\ &= \lim_{n\to\infty}\frac{1}{\pi}\int \tfrac{1}{2}\iota r\,\frac{\cos\frac{1}{2}\phi-e^{(n+\frac{1}{2})\iota\phi}}{\sin\frac{1}{2}\phi}\,d\phi.\end{aligned} \tag{34}$$

Since r cosec $\frac{1}{2}\phi$ is continuous, the term involving n vanishes in the limit, and

$$\Sigma A_n=\frac{1}{2\pi}\,\iota\int_0^{2\pi} r\cot\tfrac{1}{2}\phi\,d\phi, \tag{35}$$

whence

$$\Sigma B_n=0\,;\quad \beta=\Sigma C_n=\frac{1}{4\pi a}\int_0^{2\pi}\frac{\eta'}{1-\cos\theta}\,d\theta. \tag{36}$$

Near the trailing edge η' would vary like θ^2 if S was smooth there; but if the aerofoil has a finite angle at the trailing edge the curve has a sudden bend, and the index is a little less than 2. But this does not affect the convergence.

We have already

$$B_0 = \frac{1}{2\pi}\int_0^{2\pi}\frac{\eta'}{2a}\frac{d\phi}{\sin\phi}\ ;\ C_2 = \frac{1}{\pi a}\int_0^{2\pi}\eta'\cos\phi\,d\phi, \tag{37}$$

while

$$B_0 + B_1 - B_2 = \frac{1}{2\pi a}\int_0^{2\pi}\eta'\left(\tfrac{1}{2} + \cos\phi - \cos 2\phi\right)\operatorname{cosec}\phi\,d\phi. \tag{38}$$

It appears that β and B_0 are always positive for an aerofoil of finite thickness with the medial line everywhere on the same side of the chord; the signs of C_2 and $B_0 + B_1 - B_2$ depend on the form of the section. To make the moment coefficient vanish at zero lift we must have

$$\int_0^{2\pi}\eta'\left(2\cos\phi + \frac{1}{1-\cos\phi}\right)d\phi = 0 \tag{39}$$

4. These results can be tested by comparison with those known for the thin two-term Joukowsky aerofoil obtained by transforming the circle with radius a and centre $z_0 = r_0e^{i(\pi-\lambda)}$ according to the rule

$$\zeta' = z + c^2/z, \tag{40}$$

where a is slightly greater than c, and chosen to make the point $+c$ lie on the circle. If

$$ae^{-i\beta} = c - z_0, \tag{41}$$

we find

$$\eta' = (a - c^2/a)\sin\theta - 2r_0\sin\theta\cos(\lambda+\theta), \tag{42}$$

$$k_L = \pi\left(1 + \frac{r_0}{c}\cos\lambda\right)(\alpha+\beta), \tag{43}$$

$$h_M = \tfrac{1}{4}k_L + \tfrac{1}{4}\pi\beta. \tag{44}$$

The β determined by the methods of the present paper is the same as that involved in (41), while

$$B_0 = \frac{r_0}{a}\cos\lambda, \tag{45}$$

and $B_0 + B_1 - B_2$ and C_2 are zero. The coefficient $\tfrac{1}{4}$ in (44) is a standard result.[*]

[*] Glauert, 'Aerofoil and Airscrew Theory,' p. 86.

5. On the boundary we get

$$Ue^{-\iota\alpha} - Ue^{\iota\alpha}e^{-2\iota\phi} + 2\iota U \sin(\alpha+\beta)\, e^{-\iota\phi}$$
$$= 2\iota U e^{-\iota\phi}\,[\sin(\phi-\alpha) + \sin(\alpha+\beta)], \qquad (46)$$

$$1 - \Sigma\,(n-1)\,A_n\,(a/z)^n = 1 - \Sigma\,(n-1)\,A_n\,e^{-n\iota\phi}, \qquad (47)$$

$$1 - a^2/\zeta^2 = 1 - (1-2r)\,e^{-2\iota\theta} = e^{-\iota\theta}\{2\iota\sin\theta + 2re^{-\iota\theta}\}, \qquad (48)$$

whence by (20) the resultant velocity q is

$$U\{\sin(\phi-\alpha) + \sin(\alpha+\beta)\}\,|\,1 + \Sigma\,(n-1)\,A_n\,e^{-n\iota\phi}\,| / |\,(\sin\theta - \iota r e^{-\iota\phi})\,| \qquad (49)$$

The first variable factor has the same form as for a flat plate, apart from the presence of β. It vanishes when $\phi = -\beta$ or $\phi = \pi + 2\alpha + \beta$. The former point is the trailing edge, the latter the point of stagnation near the leading edge. At the latter

$$\theta = \phi + \Sigma(-1)^n C_n \quad \text{nearly}$$
$$= \pi + 2\alpha + 2\Sigma C_{2m} = \pi + \gamma, \quad \text{say}. \qquad (50)$$

The disappearance of the C's with odd suffixes implies that the displacement of the point from its position for a flat plate depends on the differences between the camber on the fore and aft sides of the wing. We find easily that

$$\Sigma C_{2m} = \frac{1}{4\pi a}\int_0^{2\pi} \frac{\eta' \cos\theta}{\sin^2\theta}\, d\theta.$$

Since for most wings η' is greater when $\theta > \frac{1}{2}\pi$, this sum is in general negative, and θ is displaced towards π, and the point of stagnation towards the leading edge.

The second factor never departs far from unity except perhaps close to the trailing edge. The denominator is

$$|\sin\theta - r\,(\iota\cos\theta + \sin\theta)| = \{\sin^2\theta\,(1-r)^2 + r^2\cos^2\theta\}^{\frac{1}{2}}$$
$$= \{r^2 + \sin^2\theta\,(1-2r)\}^{\frac{1}{2}}, \qquad (51)$$

which is nearly $\sin\theta$ except near the edges. At the front edge it is r, so that the curvature of the edge plays an important part in determining the velocity. In the neighbourhood we have nearly

$$q = \frac{U \sin(\phi-\alpha) + \sin(\alpha+\beta)}{\{\sin^2\theta + r^2\}^{\frac{1}{2}}}\,.$$
$$= \frac{2U\sin\theta\sin\frac{1}{2}(\pi+\gamma-\theta)}{\left\{\left(1-\frac{\xi'^2}{4a^2}\right)^2 + \frac{\eta'^2}{4a^2}\right\}^{\frac{1}{2}}} \qquad (52)$$

nearly, using (6); and when ξ' is nearly $-2a$, equal to $-2a+\tau$, we have

$$\eta'^2 = 2\rho\tau; \quad 1 - \xi'^2/4a^2 = \tau/a, \tag{53}$$

and

$$q = \frac{2\mathrm{U}\sin\frac{1}{2}(\pi+\gamma-\theta)}{\{(\frac{1}{2}\rho+\tau)/a\}^{\frac{1}{2}}}. \tag{54}$$

Close to the leading edge q is nearly $2\sqrt{2}\,(a/\rho)^{\frac{1}{2}}\,\mathrm{U}\sin\alpha$, but falls off when the distance from the edge exceeds the radius of curvature there.

The Effect of Electric Fields on the Emission of Electrons from Conductors.

By A. T. WATERMAN, Assistant Professor of Physics, Yale University, U.S., National Research Fellow, at King's College, London.

(Communicated by O. W. Richardson, F.R.S.—Received April 13, 1928.—
Revised May 16, 1928.)

§ 1. In a recent paper Houston[*] has proposed an explanation of the extraction of electrons from metals by intense electric fields, which is based upon Sommerfeld's[†] modification of the Lorentz electron theory of metals. The writer has also investigated this problem from the same standpoint, but reached the conclusion that such an explanation is inadequate, at least in its present form. Houston's paper should, however, be regarded as a treatment of the Schottky effect for very intense fields, but, as will be shown, if the Sommerfeld electron theory is accepted the expression for the Schottky effect should be modified, and this modification becomes significant at high fields.

The object of this paper then is (1) to derive an expression on the Sommerfeld theory for the thermionic current in the presence of an electric field, and (2) to point out the difficulties in the way of accepting this effect as the correct explanation of the "field currents" examined in detail by Millikan and Eyring[‡], Gossling,[§] del Rosario,[||] and others. In the course of the paper it

[*] 'Z. Physik,' vol. 47, p. 33 (1928).

[†] 'Z. Physik,' vol. 47, p. 1 (1928).

[‡] 'Phys. Rev.,' vol. 27, p. 51 (1926).

[§] 'Phil. Mag.,' vol. 1, p. 609 (1926).

[||] 'J. Franklin Inst.,' vol. 203, p. 243 (1927), and vol. 205, p. 16 (1928).

will be necessary to work out the volume distribution, *i.e.*, degree of penetration, of a surface electrical charge on the basis of the Sommerfeld electron theory.

In Houston's paper the applied electric field is considered to diminish the energy required to liberate an electron, in the manner treated by Schottky.* Sommerfeld's expression for the usual thermionic current density is

$$\mathrm{I} = \mathrm{A}k^2\mathrm{T}^2e^{-\phi/k\mathrm{T}}, \tag{1}$$

where $\phi = w - \Omega$. Ω being the "electron pressure" energy ($= \mathrm{W}_i$ in Sommerfeld's notation), and w the total energy necessary to liberate an electron from the metal ($= \mathrm{W}_a$ in Sommerfeld's notation), so that ϕ is the net energy required for liberation of an electron—the ordinary work function. Thus w is considered to be decreased by action of an accelerating field, and the current density is thereupon increased by the consequent reduction of ϕ. As Houston points out, equation (1) does not hold when $\phi \leqslant 0$, *i.e.*, when by action of the field w becomes equal to or less than Ω. In that case we have, on Sommerfeld's theory (equation (5) in Houston's paper, *loc. cit.*):

$$\mathrm{I} = \frac{\mathrm{A}}{2}\left[\phi^2 + 2k^2\mathrm{T}^2\left(\frac{\pi^2}{6} - e^{\phi/k\mathrm{T}} + \tfrac{1}{4}e^{2\phi/k\mathrm{T}} - \tfrac{1}{9}e^{3\phi/k\mathrm{T}} + \ldots\right)\right], \tag{2}$$

where ϕ is negative or zero, and A is the same constant as in equation (1). For the sake of theoretical completeness it may be pointed out that equation (1) does not hold exactly when ϕ approaches zero and is positive (w slightly $> \Omega$). The general formula for the thermionic current density when $w > \Omega$ is

$$\mathrm{I} = \mathrm{A}k^2\mathrm{T}^2\left(e^{-\phi/k\mathrm{T}} - \tfrac{1}{4}e^{-2\phi/k\mathrm{T}} + \tfrac{1}{9}e^{-3\phi/k\mathrm{T}} - \ldots\right). \tag{3}$$

This expression results, in the usual method for finding the number of electrons that come to the metal surface with a kinetic energy component (normal to the surface) greater than w, from the integral

$$\int_0^{e^{-\phi/k\mathrm{T}}} \frac{\log(1+y)}{y}\,dy,$$

on expanding $\log(1 + y)$ for $y < 1$. Equation (1) is obtained on neglecting higher order terms than the first in this expansion. Equations (2) and (3) are seen to reduce to the same value $\mathrm{A}k^2\mathrm{T}^2(1 - \frac{1}{4} + \frac{1}{9} - \ldots) = \frac{1}{12}\pi^2\mathrm{A}k^2\mathrm{T}^2$ when $\phi = 0$, as they should, while equation (1) differs from this value by the factor $\frac{1}{12}\pi^2$. The difference between equation (1) and equation (3) is thus not of great importance in numerical calculation, amounting at most to about 20 per cent. The differences may be seen by comparing the upper half of Table III with the corresponding table given by Houston (*loc. cit.*, p. 35), who uses equation (1) in this region.

---

* 'Phys. Z.,' vol. 15, p. 876 (1914), and vol. 14, p 63 (1923).

§ 2. *Modification of the Schottky Equation according to Sommerfeld's Theory.*

A characteristic of the Sommerfeld theory is that Ω depends upon the concentration n of "free" electrons at the point considered, which in this case is the emitting surface. Actually

$$\Omega = an^{2/3}, \tag{4}$$

where $a = h^2/2m \,.\, (3/4\pi G)^{2/3}$, G being the quantum weight of the spinning electron. Physically Ω represents an internal "electron pressure" energy, *i.e.*, energy possessed by an electron, on the average, which assists its escape. Since Ω is proportional to $n^{2/3}$ it should follow that the negative surface charge produced by an accelerating field would, by increasing the "electron pressure" energy Ω, facilitate electron emission, and this effect would unite with the reduction of w, as postulated by Schottky, to increase the emission current in the presence of an aiding field. The magnitude of the effect of the field upon Ω will evidently depend upon the increase of n, *at the surface of the metal*, due to the field. The solution of this problem then involves an investigation of the degree of penetration of a surface charge into the body of a conductor.

§ 3. *Volume Distribution of a Surface Electrical Charge on a Conductor.*

This matter has been investigated by J. J. Thomson[*] for the case where the internal electrons obey the classical ideal gas laws. It remains to repeat the derivation for a degenerate electron gas obeying the Fermi-Dirac statistics. The difference in the two cases is that the electron gas pressure, instead of being $p = nkT$, will be given by

$$p = bn^{5/3}, \text{ where } b = \frac{4\pi}{15}\frac{Gh^2}{m}\left(\frac{3}{4\pi G}\right)^{5/3}. \tag{5}†$$

This is the "zero-point" pressure alone, the additional term amounting to less than 0·4 of 1 per cent. of the "zero-point" pressure up to 3000° K.

Taking the yz plane as the boundary surface of the conductor with the positive x axis extending into the conductor, we have

$$dp/dx = (n_0 + \xi)\,\epsilon\, dV/dx, \tag{6}$$

[*] 'The Corpuscular Theory of Matter,' p. 81, New York (1907).

[†] *Cf.* Sommerfeld, *loc. cit.*, equation 37A.

where p = electron pressure, n_0 = normal electron concentration, ξ = concentration of electrons added in excess of normal, ϵ = electronic charge, considered as a negative number, V = potential, of conventional sign.

Also

$$p = b\,(n_0 + \xi)^{5/3}, \tag{7}$$

where b is defined by (5), and

$$\frac{d^2\mathrm{V}}{dx^2} = 4\pi\xi\epsilon. \tag{8}$$

Eliminating V and p from (6), (7) and (8)

$$\frac{d^2\,(n_0 + \xi)^{2/3}}{dx^2} = \frac{8\pi\epsilon^2}{5b}\,\xi, \tag{9}$$

which is to be solved for ξ in terms of x

Making the substitution $\mathrm{N} = (n_0 + \xi)^{1/3}$, $\mathrm{N}_0 = n_0^{1/3}$, and writing $a = 8\pi\epsilon^2/5b$, and integrating

$$2\mathrm{N}\frac{d\mathrm{N}}{dx} = \pm\,[2a\,(\tfrac{2}{5}\mathrm{N}^5 - \mathrm{N}_0^3\mathrm{N}^2 + \tfrac{3}{5}\mathrm{N}_0^5)]^{1/2}.$$

The constant of integration C being determined by the condition that when $\xi = 0$, *i.e.*, $\mathrm{N} = \mathrm{N}_0$, therefore $d\mathrm{N}/dx = 0$.

Putting $\mathrm{N} = \mathrm{N}_0 + \nu$,

$$dx = \pm\,(2/a)^{1/2}\,(\mathrm{N}_0 + \nu)\,d\nu\,(\tfrac{2}{5}\nu^3 + 2\mathrm{N}_0\nu^2 + 4\mathrm{N}_0^2\nu + 3\mathrm{N}_0^3)^{-1/2}. \tag{10}$$

For values of $\nu < 0{\cdot}3\,\mathrm{N}_0$, the term $\tfrac{2}{5}\nu^3$ is less than $0{\cdot}002$ of the entire expression under the radical and may be neglected. This condition covers values of ξ, the added electron concentration, up to $1{\cdot}2$ times n_0, the normal concentration. Then integrating

$$x = \pm\,(2/a\mathrm{N}_0)^{1/2}\{-1/\sqrt{3}\,.\log[(2\eta^2+4\eta+3)^{1/2}+2\eta/\sqrt{3}+\sqrt{3}]+1/\sqrt{3}\,.\log\eta$$
$$+1/\sqrt{2}\,.\log(\sqrt{2\eta^2+4\eta+3}+\eta\sqrt{2}+\sqrt{2})+1/\sqrt{2}\,.\log\mathrm{N}_0\}+\mathrm{B}, \tag{11}$$

where $\eta = \nu/\mathrm{N}_0$. On evaluating the expression in the brackets for various values of η, assuming N_0 of the order of magnitude required by a normal electron concentration equal to the concentration of atoms in the metal, it is found that for values of η between 10^{-3} and 10^{-1} only the terms

$$1/\sqrt{3}\,.\log\eta + 1/\sqrt{2}\,.\log\mathrm{N}_0$$

need be retained, with an accuracy of 1 per cent.* This corresponds to a range of ξ from $0 \cdot 0003\ n_0$ to $0 \cdot 33\ n_0$. Outside the range quoted this approximation holds within 2 per cent. from $\xi = 3 \times 10^{-5} n_0$ to $\xi = 2 \cdot 4\ n_0$.

---

* Though not here necessary, multiplying the retained terms by $1{\cdot}01$ increases the accuracy to about 1 in 10^4 for the range considered.

Hence

$$x = \pm (2/a\mathrm{N}_0)^{1/2} (1/\sqrt{3} \,.\, \log \eta + 1/\sqrt{2} \,.\, \log \mathrm{N}_0) + \mathrm{B}, \tag{12}$$

or

$$\eta = n_0^{-1/\sqrt{6}} e^{\pm c\,(\mathrm{B}-x)}, \tag{13}$$

where $c = (3an_0^{1/3}/2)^{1/2}$.

The exact relation between ξ and η is

$$\xi/n_0 = 3\eta + 3\eta^2 + \eta^3. \tag{14}$$

With an accuracy better than 1 per cent for $\xi < 0 \cdot 03\ n_0$, we may take as a first approximation $\xi/n_0 = 3\eta$, whence, by (13):

$$\xi = \gamma e^{\pm c\,(\mathrm{B}-x)}, \text{ where } \gamma = 3n_0^{1-1/\sqrt{6}}.$$

The constant B may now be determined from the boundary condition that $\int_0^\infty \xi\, dx = \sigma'$, where $\sigma' = \sigma/\epsilon$; σ being the surface charge. σ' is therefore the total number of electrons added per unit of surface.

It follows that

$$\xi = c\sigma' e^{-cx}. \tag{15}$$

Thus, provided the concentration of electrons added does not approach the normal concentration too closely, the excess electron density falls off exponentially with the distance from the surface, as in the classical case.

Assuming that the value of G is 2, as expected on theoretical grounds and for which Fowler* has offered experimental evidence in the case of thermionic emission, and using $n_0 = 6 \cdot 20 \times 10^{22}$ per cubic centimetre for tungsten, a good representative value, we find

$$c = (3an_0^{1/3}/2)^{1/2} = 4 \cdot 3^{1/6} \pi^{5/6} m^{1/2} h^{-1} \epsilon n_0^{1/6} = 1 \cdot 72 \times 10^8. \tag{16}$$

The maximum value of ξ is therefore $1 \cdot 72 \times 10^8\ \sigma'$, which for a surface charge of 10^{13} electrons per square centimetre ($1 \cdot 6 \times 10^{-6}$ coulombs per square centimetre) means a maximum added electron density at the surface of $1 \cdot 72 \times 10^{21}$ electrons per cubic centimetre, or about $0 \cdot 03\ n_0$. Thus for surface charges of this magnitude or less the formula (15) should apply. In Millikan and Eyring's experiment (*loc. cit.*) the maximum surface charge was of the order of 10^{12} electrons per square centimetre.

* 'Roy. Soc. Proc.,' A, vol. 118, p. 52 (1928).

The actual distribution of electron density inward from the surface requires some consideration, lest it be found that the theoretical maximum density at the surface exists in an impossibly thin layer. Table I gives computed values of ξ/σ', for different values of x, the distance from the surface.

Table I.

x.	ξ/σ'.	ξ.
Depth (cm.)	Excess electrons per c.c. / Electrons per cm. of surface.	Excess electron density* (per c.c.).
0	$1\cdot72 \times 10^{8}$	$1\cdot7 \times 10^{20}$
10^{-10}	$1\cdot69 \times 10^{8}$	$1\cdot7 \times 10^{20}$
10^{-9}	$1\cdot45 \times 10^{8}$	$1\cdot5 \times 10^{20}$
10^{-8}	$3\cdot11 \times 10^{7}$	$3\cdot1 \times 10^{19}$
5×10^{-8}	$2\cdot83 \times 10^{4}$	$2\cdot8 \times 10^{16}$
10^{-7}	$6\cdot38$	$6\cdot4 \times 10^{12}$
$2\cdot7 \times 10^{-7}$	10^{-12}	1

* For surface charge of 10^{12} electrons per square centimetre.

It is seen that the electron density changes relatively little within a layer about one atom deep, after which it falls off rapidly, becoming entirely negligible at a depth a little beyond 10^{-7} cm. The mean electron density for a surface layer 10^{-8} cm. thick comes out $8\cdot2 \times 10^{7}$ for ξ/σ', or practically half the theoretical maximum at the surface. It therefore does not seem impossible that electron densities of the order of magnitude of the theoretical maximum at the surface may in fact occur

For the number of electrons per unit of surface, σ', in terms of a field E applied to produce the surface charge:

$$\sigma' = E/4\pi e = 5\cdot55 \times 10^{5} \text{ E v./cm.}$$

Therefore the maximum value of the electron concentration added ($x = 0$) will be

$$\xi_m = c\sigma' = 9\cdot55 \times 10^{13} \text{ v./cm.} \tag{17}$$

§ 5. *Effect of Increase in Electron Surface Concentration on Electron Pressure Energy* Ω.

As in equation (4),

$$\Omega = an^{2/3} = \Omega_0 (1 + \xi/n_0)^{2/3},$$

where $\Omega_0 = an_0^{2/3}$, the normal "electron pressure" energy. When $\xi/n_0 < 1$, as here, the change in Ω due to a surface charge will be

$$\Delta\,\Omega = \tfrac{2}{3}ac\sigma' n_0^{-1/3} = 9{\cdot}71 \times 10^{-35}\ \xi_m \text{ ergs for tungsten.}$$

Substituting from (17), the change in Ω (in ergs) produced by a field of E volts/centimetres, at a tungsten surface, is given by

$$\Delta\,\Omega \text{ ergs} = \alpha \mathrm{E} \text{ v./cm.} \tag{18}$$

where $\alpha = h/(2^4 \,.\, 3^{7/6} \,.\, 5^2 \pi^{5/6} m^{1/2} n_0^{1/6}) = 9{\cdot}27 \times 10^{-21}$ for tungsten.

Thus the thermionic work function ϕ will be decreased by an amount given by (18).

The penetration of the surface charge, on any free electron theory, should give rise to a factor in the Schottky effect which has apparently been overlooked hitherto. In the present calculation, by (15), the (volume) charge density ρ at a depth x within the conductor is $\rho = c\sigma e^{-cx}$. Using Poisson's equation subject to the boundary condition $d\mathrm{V}/dx = 0$ at $x = \infty$, we have $d\mathrm{V}/dx = 4\pi\sigma e^{-cx}$, whence the potential difference between the surface ($x = 0$) and the interior ($x = \infty$) is $\Delta \mathrm{V} = 4\pi\sigma/c$. This potential difference will diminish ϕ by an amount

$$\Delta\phi \text{ (ergs)} = 4\pi\sigma\varepsilon/c = \varepsilon/300\, c \,.\, \mathrm{E} \text{ (v./cm.)} \tag{18'}$$

At first sight it might appear that the expressions (18) and (18′) should both appear in the total $\Delta\phi$ responsible for the Schottky effect.

The coefficient $\varepsilon/300\,c$ is identical with α, as previously derived in (18). Thus we have merely arrived at the same result by another method. It is interesting to note then that the effect of a surface charge is to decrease ϕ by the addition of potential energy to the internal electrons, and that this addition of potential energy naturally finds its expression on the Sommerfeld theory in the electron pressure energy Ω, as it should.

§ 6. *The Schottky Term.*

The theory devised by Schottky (*loc. cit.*) in explanation of the effect of an electric field on thermionic currents should also hold in this case. In it the potential drop corresponding to w, the energy required to liberate an electron, is affected. In the absence of space charge, the change in this potential drop by a field E is

$$\Delta w \text{ (ergs)} = -\,\beta \mathrm{E}^{1/2} \text{ v./cm.,} \quad \text{where} \quad \beta = 6{\cdot}02 \times 10^{-16}. \tag{19}$$

§ 7. *The Thermionic Current Density under an Applied Electric Field.*

In the presence of an accelerating field, E :

$$\phi = \phi_0 - \alpha E - \beta\sqrt{E}, \tag{20}$$

where $\phi_0 = w_0 - \Omega_0$ is the thermionic work function, experimentally determined, at zero field.

The ordinary thermionic current density becomes therefore

$$I = A'T^2 e^{(-\phi_0 + \alpha E + \beta\sqrt{E})/kT}, \tag{21}$$

where $A' = 2\pi m h^{-3} \epsilon k^2 = 60$ amp /cm², G being considered cancelled by a reflection coefficient of 50 per cent. (*cf.* Fowler, *loc. cit.*) ; $\alpha = 9 \cdot 27 \times 10^{-21}$; $\beta = 6 \cdot 02 \times 10^{-16}$; E in volts per centimetre.

Equations (3) and (2) should be used for $\phi \geqslant 0$ and $\phi \leqslant 0$, using the value of ϕ given by equation (20).

Equation (21) may be written in another form, viz.,

$$I = I_0 e^{(\alpha E + \beta\sqrt{E})/kT}, \tag{22}$$

where I_0 = current density for zero field at the same temperature.

This equation (22) is the revised expression for the Schottky effect, based on Sommerfeld's electron theory. It reduces to the usual Schottky effect equation if $\alpha = 0$. It is evident that the new term, αE, appearing in the exponent will become increasingly important as the electric field is increased. The Schottky coefficient β is independent of the nature of the hot metal, while the coefficient α of the new term depends upon the metal, though only slightly, α being inversely proportional to the sixth root of n_0. The extreme variation in n_0 according to Sommerfeld's interpretation (one free electron per atom) is by a factor of 10 (between Ni, Fe or Cu and the alkali metals), which would mean a variation in α of about 50 per cent.

It may be worth noting that in the classical electron theory the expression for the Schottky effect should be of the same form (for $\phi > 0$, the usual case) as the expression (equation (22)) here derived, namely,

$$I = I_0 e^{\gamma E + \beta\sqrt{E}/kT}. \tag{23}$$

The quantity γX results in exactly the same way as the second derivation (from Poisson's equation) of αX under § 5, and is likewise due to the penetration of the surface charge. The latter is of the same form under the same limitations ($\xi < n_0$), viz., $\rho = c'\sigma e^{-c'x}$, where in this case $c' = 2\epsilon\sqrt{\pi n_0/kT}$ as compared with the Sommerfeld c given by equation (16). Consequently

$$\Delta\phi = \tfrac{1}{2}\,(kT/\pi n_0)^{\frac{1}{2}}\,E \text{ e.s.u./cm.} = \gamma E \text{ v./cm., where } \gamma = 1 \cdot 10 \times 10^{-11}(T/n_0)^{\frac{1}{2}}. \tag{24}$$

γ is thus a function of the temperature, unless n_0 is proportional to the absolute temperature. It also varies more rapidly with n_0 than does α. As for order of magnitude, for a normal free electron concentration of 10^{18} (about 2×10^{-3} of the atomic concentration), $\gamma = 1\cdot91 \times 10^{-19}$ at 300° K., being thus about 20 times greater than α. At such concentrations the electron gas would no longer be in the degenerate condition but would obey the classical laws, so that this procedure is entirely legitimate, in the absence of conclusive evidence regarding the magnitude of n_0. Owing to the distinct difference between these coefficients α and γ both in magnitude and in variation with temperature and material, accurate experimental data on the Schottky effect may prove to be an important criterion for distinguishing between the two theories.

The effect of the new coefficient α, deduced on the Sommerfeld theory, may be seen in Table II, which is computed for the case of tungsten, using the values of α and β above. Comparative values of the ratio I/I_0, for a range of field from 10^2 to 10^7 volts per centimetre, are given (1) for the ordinary Schottky equation, and (2) for this equation as here modified in equation (22), at temperatures of 300°, 1100°, and 2000° K. A column is added under $T = 1100°$ showing the order of magnitude of I/I_0 expected on the classical theory, equation (23), assuming $n_0 = 10^{18}$ at this temperature. These latter values are not given beyond $E = 10^4$ v./cm., since the exponential term γE breaks down for fields greater than this, the concentration of added electrons equalling or exceeding the normal number assumed.

Table II.—I/I_0.

Field X (v./cm.).	300° K		1100° K.			2000° K.	
	Schottky equation.	Equation (22).	Schottky equation.	Equation (22).	Equation (23).	Schottky equation.	Equation (22)
10^2	1·1576	1·1576	1·0407	1·0407	1·0410	1·0222	1·0222
10^3	1·5885	1·5889	1·1345	1·1346	1·1373	1·0719	1·0719
10^4	4·321	4·331	1·4906	1·4914	1·5271	1·2455	1·2459
10^5	102·3	104·7	3·533	3·555	*	2·002	2·009
10^6	$2\cdot3 \times 10^6$	$2\cdot8 \times 10^6$	54·1	57·6	—	8·98	9·29
10^7	$1\cdot3 \times 10^{20}$	$4\cdot0 \times 10^{20}$	$3\cdot0 \times 10^5$	$4\cdot1 \times 10^5$	—	$1\cdot0 \times 10^3$	$1\cdot2 \times 10^3$

* Derived law of penetration no longer holds.

So far as the writer is aware there have been no data published on the Schottky effect over a wide range of applied field. Dushman* gives data for applied potentials from 100 to 400 volts at constant temperatures from 1470° to 2239° K., using a tungsten loop filament. These data are, however, not particularly accurate for the purpose of verifying the Schottky equation, since the values of the field are not known—merely the voltages. In fact the data do not distinguish clearly between possible laws of the forms $I = I_0e^{cV/T}$ and $I = I_0e^{c\sqrt{V}/T}$, although, of course, according to the present calculations, the exponent in V would not be expected to make itself strongly felt at the fields probably present in Dushman's experiment.

---

* 'Phys. Rev.,' vol. 25, p. 338 (1925).

§ 8. *Application to Extraction of Electrons from Cold Metals*

The most natural assumption to make in attempting an explanation of the electron currents produced from cold metals by intense electric fields is to consider the phenomenon as an extreme case of the Schottky effect, as several investigators have done. The difficulty, however, is obviously that, while the Schottky equation gives approximately the correct variation of current with applied field, the observed effect appears to be completely independent of temperature up to about 1000° K, while the Schottky equation would require the effect to be an inverse exponential function of the temperature. Houston (*loc. cit.*) tries to avoid this difficulty by postulating actual fields (due to minute irregularities on the emitting surface) many times greater than the apparent value, so that the potential drop causing the usual work function is completely cancelled or reversed by the applied field. In that event the form of the thermionic equation is altered in such a way that the current approaches values independent of the temperature, *i.e.*, approaches the form $\sqrt{I} = a\sqrt{X} + b$ (*cf.* equation (2)). This is, however, not at all the observed variation of current with applied field, and it is then necessary to suppose an increase of emitting area with rising field of a sort to produce the observed variation. It is, of course, quite likely that the emitting area does increase with the field, but if so it is not clear why the emitting area and hence the current should not also increase with the temperature.

The same reasoning applies to the application to this case of the modified Schottky equation here derived. There is this difference, however—that the situation under which equation (2) is valid ($\phi \leqslant 0$) is brought about by lower fields. In other words, the increase in " electron pressure " energy due to a strong surface charge assists the external field in overcoming the work function. Also in this region the current varies more rapidly with the field than in the case treated by Houston, which is in the right direction to fit the facts.

Table III contains calculated values of the current density for various values of ϕ in the neighbourhood of zero (the values chosen being Houston's) for temperatures 300°, 900° and 1100° K.

Table III.

ϕ.	Current density (amp./cm.^{2}).			dV/dx (v./cm.) calculated by	
(volts).	300° K.	900° K.	1100° K.	Schottky's equation.	Equation (20).
0·63	$1·51\times10^{-4}$	$1·47\times10^{4}$	$9·61\times10^{4}$	$3·73\times10^{7}$	$3·27\times10^{7}$
0·50	$1·52\times10^{-2}$	$6·86\times10^{4}$	$3·38\times10^{5}$	$4·13\times10^{7}$	$3·47\times10^{7}$
0·32	$2·25\times10$	$7·80\times10^{5}$	$2·46\times10^{6}$	$4·80\times10^{7}$	$3·98\times10^{7}$
0·13	$3·28\times10^{4}$	$8·49\times10^{6}$	$1·70\times10^{7}$	$5·51\times10^{7}$	$4·52\times10^{7}$
0·00	$4·45\times10^{6}$	$4·00\times10^{7}$	$5·97\times10^{7}$	$6·04\times10^{7}$	$4·91\times10^{7}$
−0·13	$7·92\times10^{7}$	$1·42\times10^{8}$	$1·73\times10^{8}$	$6·59\times10^{7}$	$5·31\times10^{7}$
−0·32	$4·23\times10^{8}$	$4·94\times10^{8}$	$5·32\times10^{8}$	$7·43\times10^{7}$	$5·90\times10^{7}$
−0·50	$1·05\times10^{9}$	$1·13\times10^{9}$	$1·17\times10^{9}$	$8·30\times10^{7}$	$6·53\times10^{7}$
−0·63	$1·60\times10^{9}$	$1·68\times10^{9}$	$1·72\times10^{9}$	$8·89\times10^{7}$	$6·94\times10^{7}$

The calculations are made from equations (2) and (3) and for $\phi \leqslant 0$ Houston's figures* should agree with them. For $\phi > 0$, the values differ from Houston's for the reason stated in § 1. The last column but one gives the field required to produce the corresponding change in ϕ computed on the basis of the Schottky equation, $I = I_0 \exp (\beta\sqrt{E}/kT)$, and taking the value of ϕ_0 to be 2·94 volts, the value for thoriated tungsten. Houston's estimate of the magnitude of the field required to produce currents practically independent of temperature, namely, $1·5 \times 10^8$ volts per centimetre, results from the use of the work function for pure tungsten, whereas Millikan and Eyring (*loc. cit.*), with whose data the table is compared, employed thoriated tungsten. The last column gives the required fields as computed from equation (22). It is seen that the usual Schottky equation gives fields from 15 per cent. to 30 per cent. higher than the one here derived. All the computed values are, however, at least thirty times larger than the observed fields in Millikan and Eyring's experiment (assuming smooth cylindrical filament).

While it seems not unreasonable that fields of this magnitude may conceivably be present in such experiments, there are quite serious difficulties in the way of explaining these field currents on the same basis as the Schottky effect from Houston's standpoint, with or without the modification here introduced. First and foremost there is the magnitude of the current density required for the effect to be independent of the temperature. A current density of 10^9 or even 10^8 amperes per centimetre appears impossible on several counts. In the first place, for the smallest current measured by Millikan and Eyring, about 10^{-12} amperes, a current density of 10^8 amperes per square centimetre would mean an emitting area of 10^{-20} cm.^{2}, or emission from a spot about 10^{-10} cm. in diameter. Incidentally the cooling due to this emission from such a spot would be at the rate of 7×10^{-12} cal. per second—rather a serious problem for a spot of area 10^{-20} cm.^{2} to face, amounting to a rate of about 10^9 cal. per second per square centimetre of surface.

---

* Assuming G offset by a reflection coefficient of 50 per cent. as previously explained (*v.* equation 21).

Secondly, for fields and current densities high enough to get rid of the temperature variation the current varies very little with the applied field, *e.g.*, on doubling a field of 6×10^7 v./cm. at 300° K., the current is theoretically increased only fortyfold, as contrasted with a factor of 10^5 or 10^6 as found by observation.

Thirdly, Millikan and Eyring found the effect to be independent of temperature up to 900° K., but the emission for a given field was noticeably increased on raising the temperature to 1100° K. As seen from Table III the computed effect of temperature on the current density diminishes with rising temperature for $\phi > 0$ (equation (3)), and is practically nil for $\phi \gg 0$ (equation (3)) and is practically nil for $\phi < 0$ (equation (2)). Thus if the effect is independent of temperature at 300° K. it will remain so at higher temperatures

Attention should be called to the fact that, *if the field currents are to be explained by use of equation* (2), *i.e.*, by the case where the applied field neutralises or exceeds the work function ϕ_0, *then there is no alternative except current densities of this magnitude*, regardless of the manner of variation of the work function with applied field. The latter variation only determines the value of the field effective in reducing ϕ. Therefore it does not seem at all probable that equation (2) can prove to be the correct explanation.

As a matter of fact all sorts of secondary effects would undoubtedly occur before the current could approach such densities, *e.g*, the failure of Ohm's law within the metal.* Consideration of equation (1) indicates that one cannot hope to deal with any confidence with the case where ϕ approaches zero when changes of temperature are involved. Ω and w must both to some extent at least be functions of the temperature through thermal expansion, and it becomes highly uncertain how their difference, as it approaches zero, varies with temperature.

To the writer it would seem that the best hope of success from this general viewpoint is to try to find a condition such that a change in applied potential difference too small to detect causes a change in current equal to that produced by raising the temperature several hundred degrees. Such a situation would account for the lack of stability which is commonly experienced in experiments of this sort. Thus, using the present values for α and β in equation (22), the field, for which a change in temperature from 300° to 900° K. causes the same change in current density as a change in applied field of 5 per cent., comes out to be about $6 \cdot 3 \times 10^7$ v./cm. Assuming that the applied potential difference can only be read to 5 per cent., then for fields greater than this value the effect would seem independent of temperature but quite unstable. Now, if ϕ varied more rapidly with the field, this situation would occur at much lower fields and at far more reasonable current densities. For example, if reason might be found for increasing the coefficient α one hundredfold, then a field of about $3 \cdot 5 \times 10^6$ v./cm. would produce a current density of 10^4 amp./cm.^{2} at 300° K., while the change in field required to compensate for the increase in current on raising the temperature to 900° K. would not be detected if the potential difference could be read only to 5 per cent. For higher fields this compensating change of field would fail still more of detection.

---

* *Cf.* Bridgman, 'Proc. Nat. Acad. Sci.,' vol. 7, p. 299 (1921).

In view of the undoubted fact that the emission is extremely intense at certain spots on the emitting surface, it is possible that a space charge effect should be taken into account, but the extension of the derivation to the inclusion of space charge presents difficulties.

The theory here presented applies strictly to the case of plane parallel electrodes, while the experimental data on field currents have been obtained with coaxial filament and cylinder. The application to the latter case should be legitimate however. A re-examination of Schottky's theory as it would apply to the case of discharge from points might lead to different results.

All things considered, however, it appears as though the extraction of electrons from metals by intense electric fields involves some factor not yet considered, and it is possible that it is only to be explained in some entirely different manner, such as the treatment recently presented by Richardson.*

I am glad to take this opportunity of expressing my thanks to Prof. O. W. Richardson for his interest and stimulating criticism in connection with this work.

---

* 'Roy. Soc. Proc.,' A, vol. 119, p. 531 (1928).

An Ampere Meter for measuring Alternating Currents of Very High Frequency.

By E. B. MOULLIN.

(Communicated by R. V. Southwell, F.R.S.—Received May 21, 1928.)

Introduction.

The measurement of alternating currents of very high frequency in general presents great difficulty because all existing ammeters have a frequency correction which cannot be calculated. The dynamometer ammeter of ordinary construction is very satisfactory for currents whose frequency does not exceed a few hundred cycles per second, but is unsuitable for currents whose frequency is many kilocycles because the highly inductive winding presents an enormous impedance to the current and its presence in a circuit modifies completely the conditions obtaining therein. Further, the distribution of current over the cross section of the wire and the value of the current from turn to turn will alter with the frequency, with the result that a steady current calibration becomes invalid and the correcting factor cannot be calculated or even estimated roughly. So even if its presence can be tolerated in a circuit, such an instrument can be used only for relative measurements at one frequency and its indications cannot be reduced to absolute measure.

The ammeters in general use are thermal instruments depending on the thermal expansion of a suitable element or on the production of a thermo-electric E.M.F. in a junction placed close to a wire heated by the high frequency current. In either system the resistance of the heated element depends on the frequency of the current which heats it and so a steady current calibration cannot be used indiscriminately. The necessary configuration of the heated element and its situation with respect to surrounding alternating magnetic fields usually renders impossible the calculation of its resistance. It is usual to make the heated wire of a high resistance material and with a small diameter so as to render the calibration sensibly constant up to a high value of frequency. If the calculated resistance of such a straight isolated wire at an assigned frequency has increased by only a small fraction of 1 per cent. above the steady current value, then it will seem reasonable to suppose that the calibration of the bent and unisolated heated wire is valid to at least 1 per cent. up to this frequency. For higher frequencies there will be a correction term whose value can be estimated only very roughly. To maintain the calibration valid up to a frequency of some thousand kilocycles per second the heating wire must be so fine that it will not carry a current of more than, say, 1 ampere. To measure larger currents we are faced by the problem of providing a shunt which the shunting ratio is independent of frequency. Brief consideration will show it is very difficult to arrange a group of fine parallel wires so that each has precisely the same resistance, and further that the situation of each one is the same with respect to all the others and also the remainder of the circuit. If both conditions are not fulfilled the total current will not always divide equally among the component parallel paths and the calibration curve will be subject to a frequency correction. A common and successful method is to arrange the parallel wires as generators of a cylinder, double cone, or hyperboloid and to allow each wire to heat one of a group of thermocouples connected in series electrically.

Another method of shunting a fine wire thermal ammeter is by means of a current transformer with a closed iron core. This method has recently been developed successfully, and it is now possible to measure currents up to about 1000 amperes in value and having a frequency up to a few hundred kilocycles with an accuracy which probably is considerable.

Very recent developments in radio communication have more or less suddenly extended the important spectrum of electric currents from 1000 kilocycles to 30,000 kilocycles ($\lambda = 10$ m.), and over this extended range of frequency the accuracy of ammeters at present is very doubtful.

The author has attempted to produce an ammeter of a form which can carry, unshunted, currents up to several hundred amperes and in which the frequency error can always be calculated. He will now describe the conception, analysis and constructional development of this instrument.

(2) *Elementary System of the Ammeter.*

Consider two long straight solid circular cylinders of diameter d placed parallel to one another at a distance apart D, and let them carry the same current oppositely directed, as shown in section in fig. 1.

The two cylinders will repel one another, and if the currents i are distributed uniformly over the cross section, the repulsive force F per unit length is given by the equation $F = 2i^2/D$. If the current is alternating and of root mean square value I, the mean value of the repulsive force is given by $F = 2I^2/D$.

If we can measure the mean force we have a measure of the mean square current flowing, and the system is an ammeter which can be calibrated by steady current and used to measure alternating current. A convenient method of measuring the force is to observe the spread of the distance D when the movement of the cylinders is resisted by an elastic constraint. Let one cylinder be fixed, and let a force F move the other wire through a distance x against an elastic constraint of value μ per unit displacement. Then

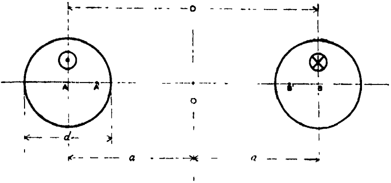


FIG. 1.

$$\mu x = 2\text{I}^2/(\text{D} + x), \text{ therefore } \mu x^2 + \mu \text{D}x - 2\text{I}^2 = 0, \text{ therefore}$$

$$x = \tfrac{1}{2}\{\sqrt{\text{D}^2 + 8\text{I}^2/\mu} - \text{D}\}.$$

The relation between x and I is a hyperbola which will not differ much from the parabola $x = 2\text{I}^2/\mu\text{D}$, if x is always small compared with D.

In the mechanical development of this instrument clearly it would be unsuitable to observe the movement of the cylinder by unaided vision or sufficiently to magnify the movement by a lever system. The movement should be observed by a high power optical microscope, and then it is convenient to restrict the greatest value of x to a small fraction of a millimetre. If the cylinders are thin wires it will be desirable to make D several millimetres and then the calibration curve will tend to be parabolic.

This ammeter will have a frequency error because the alternating current will not distribute itself uniformly over the cross section, and because the capacity between the two cylinders will cause the current to alter along the length of the parallel system. In reality the system will be complicated still further by the magnetic and electrostatic relationship of the system to external objects and fields. We shall consider each cause of frequency error separately and in the order stated above.

(3) *Change in Force due to Ununiform Current Distribution.*

When the current is alternating the distribution will depart from uniformity and from axial symmetry, but since two long parallel equal circular cylinders form a system amenable to exact mathematical treatment, the alteration of force may be calculated. It is illustrative to regard the changing distribution of current as arising from two causes. A distribution having axial symmetry would arise from the magnetic field within the wires and tend ultimately to a tubular current: this distribution having axial symmetry will not alter the force, which remains the same as if all the current was concentrated along the axes of the cylinders. But the external field of one wire will penetrate the other wire, and generate therein an eddy current system having symmetry about the diameter perpendicular to the line of centres and directed at every instant so as to tend to neutralise the penetrating field. Therefore the eddy current in the hemicylinder nearer the other wire will tend to have always the same direction as the main current in that wire, and the additional repulsive force arising therefrom will more than counterbalance the attraction from the opposite eddy current in the further hemicylinder.

We may obtain the expression for the force in the limiting condition, where the frequency tends to infinity, in a very simple manner. For in the limiting condition, there will be no flux inside the wire and the whole current will be concentrated on the surface: and in these circumstances the circumference of each wire must coincide with a line of force of the resultant field.

Now it follows readily that equal circles centred at A and B, fig. 1, will be lines of force for equal and oppositely directed currents centred at A′ and B′, where A′ and B′ are mutually inverse points for both circles. Hence it follows that these currents produce a field outside these circles which is similar to that produced by the surface distribution we have postulated. Hence ultimately we have

$$F = 2I^2/A'B'.$$

It can be shown that $A'B' = \sqrt{D^2 - d^2}$,
therefore

$$F = 2I^2/\sqrt{D^2 - d^2} \quad \text{in the limit.} \tag{1}$$

We may notice that the ultimate increase of force is small so long as D is considerably larger than d.

An expression for the force F at any frequency n may be derived. Writing $z^2 = \pi p d^2/\rho$, where $p = 2\pi n$ and ρ is the specific resistance of the material, it may be shown that

$$F = \frac{2I^2}{D}\left\{1 - \tfrac{1}{2}\frac{d^2}{D^2}\phi_1 + \tfrac{1}{8}\frac{d^4}{D^4}(2\phi_1^2 - \phi_2) + \ldots\right\},$$

where roughly $-\phi_1 = z^4/48$ and $-\phi_2 = z^4/384$ for values of $z < 1{\cdot}5$ and $-\phi_1 = 1 - \sqrt{2/z}$ and $-\phi_2 = 1 - 2\sqrt{2/z}$ within 5 per cent. for values of $z > 3{\cdot}5$.*

Thus

$$\mathrm{F} = \frac{2\mathrm{I}^2}{\mathrm{D}}\left\{1 + \frac{1}{96}\frac{d^2}{\mathrm{D}^2}z^4 + \frac{z^4}{3072}\frac{d^4}{\mathrm{D}^4} + \ldots\right\} \text{ with } z < 1{\cdot}5,$$

and

$$\mathrm{F} = \frac{2\mathrm{I}^2}{\mathrm{D}}\left\{1 + \tfrac{1}{2}\frac{d^2}{\mathrm{D}^2} + \frac{3}{8}\frac{d^4}{\mathrm{D}^4} - \frac{1}{\sqrt{2z}}\frac{d^2}{\mathrm{D}^2}\left(1 + \frac{d^2}{\mathrm{D}^2}\right) + \ldots\right\} \text{ with } z > 3{\cdot}5. \qquad (2)$$

Formula (2) shows that the limiting value of F is approached very quickly so long as D/*d* is appreciably greater than unity. In Table I below are shown collected the value of F_0 and F_∞ and the value of F when $z = 10$, for various values of D/*d*.

Table I.

D/*d*.	F $z = 0$.	F $z = 10$.	F $z = \infty$.	$\mathrm{F}_\infty/\mathrm{F}_{10}$.
3/2	2 I^2/D	1·296 F_0	1·34 F_0	1·032
2	2 I^2/D	1·13 F_0	1·15 F_0	1·02
3	2 I^2/D	1·0476 F_0	1·06 F_0	1·01
4	2 I^2/D	1·026 F_0	1·031 F_0	1·005

We may see from this table that so long as D/*d* is greater than 3 the limiting increase of F is less than 6 per cent, and that in all cases shown the limiting value has sensibly been reached when $z = 10$.

For small current ammeters it is desirable to use thin wires of copper or tubes of aluminium spaced well apart and then the correcting factor is small in value and does not become appreciable until the frequency is very high. For example, Table II shows the correcting factor at various frequencies for copper wires 1·1 mm. diameter and placed 4 mm. apart.

Table II.

Wave-length λ or frequency n.	$\lambda = 10^4\ m$ 3×10^4.	$\lambda = 3000\ m$ 10^5	$\lambda = 600\ m$ 5×10^5	$\lambda = 300\ m$ 10^6.	$\lambda = 30\ m$ 10^7.
Correcting factor for copper wires	1·012	1·024	1·031	1·034	1·04

---

* Values of ϕ_1 and ϕ_2 may be found from 'Phil. Mag.,' vol. 19, p. 49 (1910), and 'Phil. Trans.,' A, vol. 222, p. 64 (1921).

The author is indebted to Mr. R. R. M. Mallock for these expressions.

Difficulties of mechanical construction make it desirable to replace at least one cylinder by a thin-walled tube. An exact expression for the force between hollow tubes or between a tube and a cylinder has not yet been derived because the analysis is extremely cumbersome and laborious. But the limiting value of the force cannot be affected by the absence of the solid core, so that for very high frequencies expression (1) is applicable to a tube. But it still remains to investigate the rate of approach to the limiting value. Consider a thin tube of mean radius R and thickness t and specific resistance ρ, which is in the field of a line current $I \sin pt$ situated at B as in fig. 2. Then if the field inside the cylinder due to the eddy currents in the cylinder be ignored, it follows that the eddy currents induced by the current at B will be in phase quadrature with this current, and then the mean force on the cylinder will be zero. If AB is large compared with R, the force on the cylinder will be affected more by change of phase of the eddy currents than by their change of distribution round the cylinder. Consequently the mean value of the force must depend largely on a factor of the form $p^2R^2(p^2R^2 + \alpha\rho^2/t^2)$.

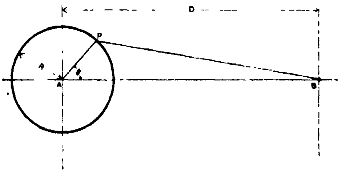


FIG. 2.

[*Note.*—It follows from elementary analysis that the current density σ is given by

$$\sigma = \frac{2pI}{\rho} \sum_{1}^{\infty} \frac{1}{n}\left(\frac{R}{D}\right)^n \cos n\theta \cos pt, \text{ for very low frequencies,}$$

and by

$$\sigma = \frac{I}{\pi R} \sum_{1}^{\infty} \left(\frac{R}{D}\right)^n \cos n\theta \sin pt, \text{ for infinite frequency.}$$]

Let ϕ_0 be the potential outside the cylinder of the eddy currents in the cylinder. Then it may be shown that*

$$\phi_0 = 2m\mathrm{I} \sum_1^\infty \left(\frac{\mathrm{R}^2}{\mathrm{D}r}\right)^n \frac{n \cos pt + m \sin pt}{n\,(m^2 + n^2)} \,.\, \sin n\theta,$$

where $m = 2\pi pt\mathrm{R}/\rho$.

Whence it follows that H the field at B is given by

$$\frac{1}{\mathrm{D}}\left(\frac{\partial \phi_0}{\partial \theta}\right)_{\theta=0} = \mathrm{H} = \frac{2m\mathrm{I}}{\mathrm{D}} \sum_1^\infty \left(\frac{\mathrm{R}^2}{\mathrm{D}^2}\right)^n \frac{n \cos pt + m \sin pt}{(m^2 + n^2)},$$

$$\int_0^{2\pi/p} \frac{\mathrm{HI} \sin pt \,.\, dt}{2\pi/p} = \frac{\mathrm{I}^2}{\mathrm{D}} \sum_1^\infty \left(\frac{\mathrm{R}^2}{\mathrm{D}^2}\right)^n \frac{m^2}{m^2 + n^2}.$$

If the tube carries a current I sin pt, then the total force is

$$\mathrm{F} = \frac{\mathrm{I}^2}{\mathrm{D}}\left\{1 + \frac{m^2}{1 + m^2}\frac{\mathrm{R}^2}{\mathrm{D}^2} + \frac{m^2}{4 + m^2}\frac{\mathrm{R}^4}{\mathrm{D}^4}\right\} \text{ approx. if } \mathrm{D} >> \mathrm{R}.$$

This expression gives the force between a line current and a tube, but needs further consideration before it may be applied to calculate the force between a solid cylinder and a tube. We will consider only frequencies which make z for the solid cylinder greater than, say, 10. Then the surface of the copper cylinder cannot differ appreciably from a line of force, and the field of all the current in it can differ inappreciably from that of a line current centred at an inverse point such as B′ in fig. 1. Therefore in fig. 2 the solid cylinder must be supposed centred at a point O on AB produced such that OB . OA = R².

We may best explain the process of calculating the force between a solid cylinder and a thin tube by means of a numerical example which refers to a specific instrument to be described later.

The solid cylinder is of copper 1·1 mm. diameter, and the thin tube is of aluminium 1·1 mm. external diameter and 0·025 mm. thick. When the frequency is 10^6 cycles, z for the solid copper cylinder is 11·8, and so for all greater frequencies we shall suppose the current in it may be considered concentrated at the point B′ of fig. 1. The cylinders are mounted so that their centre line distance is 4·0 mm. Hence the distance D in fig. 2 is approximately 4 $(1 - \mathrm{R}^2/\mathrm{D}^2) = 4 \times 0{\cdot}98$ mm., whence $\mathrm{R}^2/\mathrm{D}^2 = 1/50$. For the aluminium tube of mean radius 0·537 mm., we find that $m = 1{\cdot}79n/10^6$. Hence may be calculated the following table :—

---

* See 'Camb. Phil. Soc. Proc.,' vol. 23, p. 902 (1927), F. W. Carter.

Table III.

n.	5×10^5.	10^6.	2×10^6.	3×10^6.	4×10^6.	Infinite.
F_n/F_0	1·024	1·035	1·038	1·039	1·04	1·04

On comparing Table II and Table III we see that the one tubular conductor retards slightly the attainment of the limiting correcting factor, but to the accuracy of the calculation the two tables do not differ for frequencies above 1000 kilocycles, and then we may consider the limiting correcting factor of 4 per cent. has sensibly been reached.

We may approach the tube problem in another way. When the frequency is very high the current density decreases inwards from the surface of a solid in a manner which is sensibly exponential, and therefore the current may be said sensibly to be concentrated in a surface skin of finite thickness. If the depth of this skin is small compared with the radius of the cylinder, then the rate of penetration will differ insensibly from that appropriate to an infinite flat plate. If the density at the surface of such a plate is σ_0, then it is well known that $\sigma = \sigma_0 \varepsilon^{-mx} \cos (pt - mx)$, where $m^2 = 4\pi^2 n/\rho$. Therefore at a depth $4/m$ from the surface the current density will have fallen to about 2 per cent. of its surface value, and so we say that the depth of penetration is about $4/m$. Thus if t is the thickness of a tube, we may reasonably suppose that the limiting value of the force has sensibly been reached when $t = 4/m$, which is when $z = 2\sqrt{2}\, d/t$. If we apply this criterion to the aluminium tube specified previously we conclude that the limiting force has been reached when $n = 10^8$, but our more detailed work has shown it has been reached much sooner, namely, when $n = 3 \times 10^6$. However, the simple criterion $z = 2\sqrt{2}\, d/t$ would find useful application to thick tubes.

We have now found how to calculate the change of force due to non-uniform current density round the cylinders. For the value of D/d contemplated it is always less than 4 per cent.

We will now investigate the change of force which arises from the current distribution along the parallel wire system.

(4) *Change of Calibration due to Current Distribution along the Parallel Wire System.*

We are contemplating two long parallel wires at a distance apart which is small compared with their length; for example, two wires about 4 mm. apart and each having a length of from 100 to 150 mm. We may apply with confidence to this system the analysis appropriate to a long uniform telegraph line and, following Kelvin, Heaviside and others, suppose the capacity and inductance per unit length are constant. In fact, we suppose the distribution of electric lines along any elementary section differs inappreciably from the distribution which would obtain if the whole of each cylinder were an equipotential surface, and we suppose the distribution of flux threading any elementary section differs inappreciably from that which would obtain if the current were the same at each point of the system. The system is depicted diagrammatically in fig. 3. Let the capacity and inductance per unit length be respectively C and L, and let the potential of and the current at a point on the upper wire distant x, from the origin A, be v and i respectively. If the system is isolated in free space, then from symmetry, these quantities will be $-v$ and $-i$ at the similar point on the lower wire. It is well known that the current distribution is little affected by resistance except when the frequency is close to a natural frequency of the system. The complete analysis is well known, and since we do not contemplate using the system near a natural frequency we will not complicate and obscure the main issue by taking account of resistance whose presence can make only a second order correction.

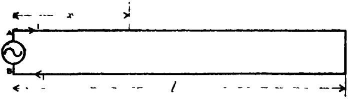


FIG. 3.

It follows readily that

$$-2\frac{\partial v}{\partial x} = L\frac{\partial i}{\partial t} \quad \text{and} \quad -\frac{\partial i}{\partial x} = 2C\frac{\partial v}{\partial t},$$

therefore

$$\frac{\partial^2 i}{\partial x^2} = LC\frac{\partial^2 i}{\partial t^2} \equiv \alpha^2\frac{\partial^2 i}{\partial t^2}. \tag{3}$$

Let the current entering at A be $-\mathrm{I}\cos pt$ and leaving at B be $\mathrm{I}\cos pt$.

When $x = l$ …, $v = 0$, therefore $\partial i/\partial x = 0$.

Whence

$$i = -\mathrm{I}\,(\cos p\alpha x + \tan p\alpha l\,.\,\sin p\alpha x)\cos pt.$$

$$= -\mathrm{I}\,\frac{\cos p\alpha\,(l - x)}{\cos p\alpha l}\,.\,\cos pt. \tag{4}$$

Similarly it follows that

$$v = \frac{I}{2}\sqrt{\frac{L}{C}}\,\frac{\sin p\alpha(l-x)}{\cos p\alpha l}\,.\,\sin pt. \tag{5}$$

Let the surrounding medium have unit permeability and dielectric constant k, and let c be the ratio of the electromagnetic to the electrostatic unit of current or charge. Then for a parallel cylindrical system we have $LC = k/c^2$.

The force per unit length between the cylinders is a repulsion of value $2i^2/\sqrt{D^2-d^2}$ due to the magnetic field and also an attraction of value $2q^2/k\sqrt{D^2-d^2}$ due to the electrostatic field.

Hence the electrostatic attraction F_e is given by

$$F_e = \frac{2\,.\,4C^2v^2c^2}{k\sqrt{D^2-d^2}} = \frac{2I^2}{\sqrt{D^2-d^2}}\,\frac{\sin^2 p\alpha(l-x)}{\cos^2 p\alpha l}\,.\,\sin^2 pt.$$

Hence we have for the mean total force

$$F = \frac{I^2}{\sqrt{D^2-d^2}}\int_0^l \frac{\cos^2 p\alpha(l-x)-\sin^2 p\alpha(l-x)}{\cos^2 p\alpha l}\,\delta x$$

$$= \frac{I^2}{\sqrt{D^2-d^2}}\,\frac{\sin 2p\alpha l}{2p\alpha\cos^2 p\alpha l} = \frac{I^2 l}{\sqrt{D^2-d^2}}\,\frac{\tan p\alpha l}{p\alpha l} \tag{6}$$

$$= \frac{I^2 l}{\sqrt{D^2-d^2}}\left(1+\tfrac{1}{3}p^2\alpha^2l^2+\tfrac{2}{15}p^4\alpha^4l^4\right) \tag{6A}$$

approx., if $p\alpha l << 1$.

If $p\alpha l = \pi/2$, there is a current node at A and B and the entering current is zero if the resistance is zero; the force expressed in terms of the entering current must then clearly be infinite, as equation (6) shows. When $p\alpha l = \pi/4$ and $\pi/6$ respectively we find by evaluating (6) that the total force is respectively 28 per cent. and 10 per cent. greater than for the same entering current of zero frequency. Equation (6A) may be interpreted better in terms of wave-length than of frequency. When $\pi\alpha l = \pi/2$, the length of the system equals a quarter wave-length of current; since the system is immersed in a dielectric of constant k the length of the current waves on the system is thereby made $1/k$ of the length they would have in a dielectric of constant unity. Let the wave-length in free space, corresponding to a frequency $n = p/2\pi$, be λ. Then we may rewrite (6A) as follows:—

$$F = \frac{I^2 l}{\sqrt{D^2-d^2}}\left(1+\tfrac{4}{3}\,\frac{\pi^2k^2l^2}{\lambda^2}+\tfrac{32}{15}\,\frac{\pi^4k^4l^4}{\lambda^4}\right). \tag{6B}$$

We are contemplating here a system whose length is about 15 cm. and immersed in a dielectric of constant 2. Then from (6A) the force will be 2·2 per cent. greater than for zero frequency when the wave-length is 7·5 m., and the frequency is 40 million cycles per second.

We may emphasise in passing that the correction for longitudinal current distribution does not depend on the capacity of the system except in respect of the dielectric constant. Therefore it is advantageous to place the cylinders close together so as to reduce the inductance of the system, and this arrangement also has the advantage of yielding a large force for a given current.

Though we have presumed the instrument is to be used to measure currents whose wave-length is much longer than the system, yet it is still applicable when the length of the system is many wave-lengths of current. To obtain the proper correcting factor it is then advisable to include the resistance term in the analysis because attenuation may be sufficient to reduce appreciably the amplitude of successive current loops.

(5) *The Effect of External Surrounding Objects and Fields.*

Hitherto we have supposed the system isolated in space and far from all other conducting bodies or alternating fields; of course such a condition is unrealisable. The magnetic field of the loop will induce currents in neighbouring conductors, and these currents will modify the force between the cylinders. Also the cylinders will be situated in the magnetic field of the remainder of the circuit and this again will modify the force. Further, each cylinder will have capacity to earth and to other neighbouring bodies, and the conditions imposed in section (4) will be violated. All these circumstances would produce errors whose magnitude could not be calculated, and uncertainty as to their importance would reduce largely the value of the instrument. Since the effect of external conditions cannot be prevented it is necessary to impose external conditions whose effect is constant and calculable precisely. This can be done by enclosing the cylinders inside a long thick copper tube and placing them symmetrically and parallel to the axis as shown in fig. 4. The magnetic field of the two cylinders will then induce eddy currents in the thick copper shield tube which will have a value in the limit such that there is no external resultant field at any point, and then the cylinders cannot induce currents in conductors outside the shield tube. Similarly external magnetic fields will induce, in the shield tube, eddy currents whose value in the limit is such that there is no resultant field at any point inside the tube, and then the force between the cylinders is not altered by external magnetic fields. The eddy currents which the cylinders induce in the shield tube will modify the force between the cylinders, but we may calculate the limiting value (which is approached rapidly at high frequencies) of this effect very readily, provided we make R so large compared with d that we may ignore the non-uniform distribution of current round the cylinders and regard them simply as line currents. The surface distribution of current on the inside surface of the screen tube must be such that there is no external resultant field : in this condition the circumference of the screen tube is a bounding line of force of the internal field. Now the circle of radius R will be a line of magnetic force for equal and opposite line currents situated at B, and at B′ the inverse of B ; similarly for currents at A, and at A′ the inverse of A. Hence the field within the cylinder is the same as that produced by currents + I at A and − I at B, together with image currents − I at A′ and + I at B′.

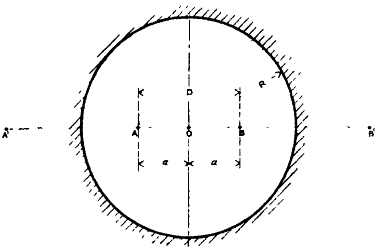


FIG. 4.

The resultant field H at B is given by

$$H = 2I\,(1/D - 1/A'B - 1/B'B).$$

Hence

$$F = \frac{2I^2}{D}\left(1 - \frac{4\,OA\,.\,OA'}{A'B\,.\,B'B}\right) = \frac{2I^2}{D}\left\{1 - \frac{4R^2}{(OA + OA')(OA' - OA)}\right\}$$

$$= \frac{2I^2}{D}\left\{1 - \frac{4R^2a^2}{R^4 - a^4}\right\}. \qquad (7)$$

Thus the effect of the screen is to reduce the force between the cylinders, and therefore tends to counteract the increase due to the non-uniform distribution round the cylinders.

We will now investigate the rate of approach to the limiting value of the effect of the screen on the force between the cylinders, and also on the screening effect from external fields. If ϕ_1 is the potential within the screen tube, of thickness t, of the eddy currents in the screen, it may be shown (see 'Camb. Phil. Soc.,' Carter, *ibid.*) that

$$\phi_1 = -4m\mathrm{I}\sum_1^\infty\left(\frac{ar}{\mathrm{R}^2}\right)^n\frac{n\cos pt + m\sin pt}{n\,(m^2+n)^2}\sin n\theta,$$

$$\text{where } n \text{ is odd and } m = \frac{2\pi tp\mathrm{R}}{\rho},$$

therefore

$$\frac{1}{a}\left(\frac{\partial\phi_1}{\partial\theta}\right)_{\substack{\theta=0\\ r=a}} = \mathrm{H} = -\frac{4m\mathrm{I}}{a}\sum_1^\infty\left(\frac{a^2}{\mathrm{R}^2}\right)^n\frac{n\cos pt + m\sin pt}{m^2+n^2},$$

therefore

$$\text{mean F} = -\frac{4\mathrm{I}}{\mathrm{D}}\sum_1^\infty\left(\frac{a^2}{\mathrm{R}^2}\right)^n\frac{m^2}{m^2+n^2}.$$

In a particular instrument to be described later, the screen tube is of copper 40 mm. radius and 3 mm. thick, and $\mathrm{D} = 2a = 4$ mm. Whence we find that $m = 0\cdot03\,n$. Therefore

$$\mathrm{F} = \frac{\mathrm{I}^2}{\sqrt{\mathrm{D}^2-d^2}}\left\{1 - \frac{m^2}{m^2+p^2}\times\frac{1}{100}\right\}\text{ approx.}$$

Hence the limiting reduction of force is $1\cdot0$ per cent., and has sensibly been attained when $m = 4$ which is when $n = 135$ cycles per second. If the screen is situated in a uniform alternating field H sin pt, then it may be shown that the eddy currents produce within the screen a field given by

$$\mathrm{H}_t = -\frac{m^2}{m^2+1}\left(\sin pt + \frac{1}{m}\cos pt\right).$$

The penetrating field will arise from the inductance coils completing the ammeter circuit, and will be in phase with the current in the cylinders forming the ammeter; therefore the quadrature term of the eddy current field will produce no net force on the cylinders and so the net penetrating field is in effect reduced by the screen in the ratio $1/(1+m^2)$.

We will now consider the capacity effect of the screen tube. Since the cylinders are enclosed completely inside an earthed screen they can have no capacity to external objects nor can external fields penetrate the screen. First suppose the far end of the loop is connected to the earthed screen tube as shown in fig. 5; and that the potential of A is V sin pt and that of B is $-$V sin pt. With this symmetrical arrangement we have not altered the system of fig. 3

except slightly to increase the capacity and decrease the inductance by corresponding amounts which are readily calculable. The longitudinal current distribution will be unaltered, and as before the total force will not be increased by 1 per cent. so long as the wave-length is greater than $50l$ But since A and B are connected to a circuit which inevitably will have large earth capacity it would be impossible to arrange that the potentials of A and B were always equally above and below the screen potential. If the ammeter is used to measure high frequency currents it will be disposed at a point in the circuit such that the potential of either A or B is zero. We will suppose that of B is zero, and then the current entering at A will leave by the point D, and the wire BD will carry a current only because a current is induced in it by the magnetic field of the current in AD. We will suppose the resistance of the circuit through BD and the case may be ignored, then the whole of the wire BD is at zero potential and must therefore carry sufficient current to make the flux zero which passes between BD and the screen. Let a flux L per unit length pass between wire AD and the screen when AD carries unit current which returns by the case and is disposed so as to produce no external field, and let a portion M of this flux pass between wire BD and the case. Then since the wires are placed symmetrically on each side of the axis, we have at every instant

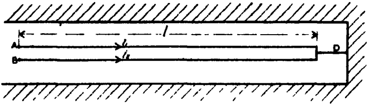


Fig. 5.

$$i_1 M = -i_2 L.$$

The flux ϕ passing between wire AD and the screen is given by

$$\phi = i_1 L + i_2 M = i_1 (L^2 - M^2)/L.$$

Hence the wire AD and the case, in the presence of the wire BD, behave as a uniform line having inductance $(L^2 - M^2)/L$ per unit length and whose capacity is the reciprocal of this expression.

It is well known that

$$L = 2 \log_e \frac{R}{r}\left(1 - \frac{a^2}{R^2 - r^2}\right) \text{ approx., if } R >> a \text{ and } R >> r,$$

and it follows readily that

$$M = 2\log_e \frac{(R + 2a)(R^2 + a^2)}{2a(R^2 + 2a^2)}$$

$$= 2\log_e \frac{R}{2a} \text{ approx., if } R >> a \text{ and } R >> r.$$

Therefore

$$\frac{L}{M} = \frac{\log_e (R/r)}{\log_e (R/2a)}.$$

In the example considered $R = 38$, $r = \frac{1}{2}$, and $a = 2$.

Whence $L/M = 2$, and $i_2 = \frac{1}{2}i_1$ and $F = 2i_1 i_1/D = I_a^2/D$.

Hence if the point B becomes in effect connected to the case by virtue of the large earth capacity associated with the point B, then the force would be halved in the limiting condition for the example considered. Therefore it is quite unsafe to attach the point D to the screen tube and the only safe method is permanently to connect B to the screen. But now corresponding points on the two cylinders will have different potentials with respect to the screen and also the current will not be the same at corresponding points; the problem is complicated and must be solved by full analysis.

(6) *Analysis of Two Cylinders placed Symmetrically within an Earthed Cylindrical Screen.*

Consider the system depicted in fig. 6 where the potential and current at a point A, distant x from the origin, on the lower wire are v_1 and i_1 respectively, and at the corresponding point on the upper wire are v_2 and i_2. Consider the

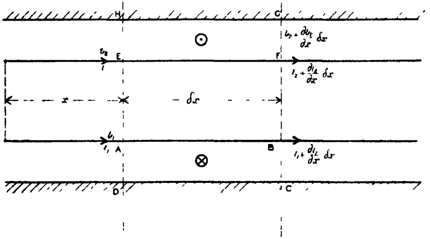

H
G
$v_2 + \frac{\partial v_2}{\partial x}\delta x$
v_2
E
F
i
$i_2 + \frac{\partial i_2}{\partial x}\delta x$
x
δx
v_1
i_1
A
B
$i_1 + \frac{\partial i_1}{\partial x}\delta x$
D
C

FIG. 6.

flux contained between two transverse planes perpendicular to the axis and placed a distance δx apart and having traces HEAD and GFBC in the figure. Let K_{11}, K_{12}, K_{22} and K_{21} be the capacity coefficients per unit length for the cylinders. Then $K_{11} = K_{22}$ from symmetry and $K_{21} = K_{12}$ from general properties of these coefficients. Let $q_1 \delta x$ be the charge on AB at time t and $q_2 \delta x$ be the corresponding charge on EF.

Then

$$q_1 \delta x = (K_{11} v_1 + K_{12} v_2)\, \delta x \quad \text{and} \quad q_2 \delta x = (K_{22} v_2 + K_{21} v_1)\, \delta x.$$

Therefore

$$-\frac{\partial i_1}{\partial x} = K_{11} \frac{\partial v_1}{\partial t} + K_{12} \frac{\partial v_2}{\partial t} \equiv C \frac{\partial v_1}{\partial t} + K \frac{\partial v_2}{\partial t},$$

and

$$-\frac{\partial i_2}{\partial x} = K_{11} \frac{\partial v_2}{\partial t} + K_{12} \frac{\partial v_1}{\partial t} \equiv C \frac{\partial v_2}{\partial t} + K \frac{\partial v_1}{\partial t}.$$

The flux through ABCD $= +(Li_1 + Mi_2)\, \delta x$, where L and M have been defined in the previous section. Therefore

$$\text{E.M.F. round ABCD} = -\left(L \frac{\partial i_1}{\partial t} + M \frac{\partial i_2}{\partial t}\right) \delta x,$$

Therefore

$$-\frac{\partial v_1}{\partial x} = L \frac{\partial i_1}{\partial t} + M \frac{\partial i_2}{\partial t}.$$

The flux through EFGH $= -(Li_2 + Mi_1)\, \delta x$,

$$\text{E.M.F. round HGFE} = -\delta v_2 = -\frac{\partial}{\partial t}\{-(Li_2 + Mi_1)\, \delta x\},$$

therefore

$$-\frac{\partial v_2}{\partial x} = L \frac{\partial i_2}{\partial t} + M \frac{\partial i_1}{\partial t}.$$

A cross relationship evidently must exist between L, M, C and K since in this system the electric and magnetic lines are everywhere orthogonal: at this stage it is important to discover what this relation is.

We will calculate first the value of C and K. If $v_1 = v_2$ then $q_1 = q_2 = q$ and $q = (C + K)\, v$. Consider the system of line charges shown in fig. 7 where

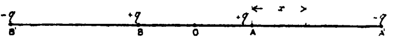


FIG. 7.

$OA \cdot OA' = OB \cdot OB' = R^2$, then these charges will make the screen tube of radius R an equipotential

$$F_x = 2q\left(\frac{1}{x} + \frac{1}{AB + x} + \frac{1}{AA' - x} - \frac{1}{AB' + x}\right),$$

therefore

$$v = 2q\int_r^R\left(\frac{1}{x} + \frac{1}{AB + x} + \frac{1}{AA' - x} - \frac{1}{AB' + x}\right)dx \text{ and } q = (C + K)\,v.$$

If $v_1 = -v_2$, then $q_1 = -q_2$ and $q = (C - K)\,v$. This system would be represented in fig. 7 by interchanging the signs of the charges at B and B′.

Then

$$v = 2q\int_r^R\left(\frac{1}{x} + \frac{1}{AB' + x} + \frac{1}{AA' - x} - \frac{1}{AB + x}\right)dx \text{ and } q = (C - K)\,v.$$

Whence, by addition,

$$\frac{2C}{2K} = \frac{\int_r^R\left(\frac{1}{x} + \frac{1}{AA' - x}\right)dx}{\int_r^R\left(\frac{1}{AB' + x} - \frac{1}{AB + x}\right)dx}.$$

To calculate L and M we must place currents + I and − I at A and B with image currents at A′ and B′.

It then follows at once that

$$L = 2\int_r^R\left(\frac{1}{x} + \frac{1}{AA' - x}\right)dx$$

and

$$M = 2\int_r^R\left(\frac{1}{AB + x} - \frac{1}{AB' + x}\right)dx,$$

therefore

$C/K = -L/M$, which is the relation we are seeking. (8)

Collecting the five equations we have, as follows :—

$$-\frac{\partial i_1}{\partial x} = C\frac{\partial v_1}{\partial t} + K\frac{\partial v_2}{\partial t}, \qquad -\frac{\partial i_2}{\partial x} = C\frac{\partial v_2}{\partial t} + K\frac{\partial v_1}{\partial t}, \tag{9}$$

$$-\frac{\partial v_1}{\partial x} = L\frac{\partial i_1}{\partial t} + M\frac{\partial i_2}{\partial t}, \qquad -\frac{\partial v_2}{\partial x} = L\frac{\partial i_2}{\partial t} + M\frac{\partial i_1}{\partial t}, \tag{10}$$

$$-C/K = L/M. \tag{8}$$

Whence we get from (9) that

$$\frac{K}{C}\frac{\partial i_2}{\partial x}-\frac{\partial i_1}{\partial x}=\frac{C^2-K^2}{C}\frac{\partial v_1}{\partial t}, \qquad \frac{K}{C}\frac{\partial i_1}{\partial x}-\frac{\partial i_2}{\partial x}=\frac{C^2-K^2}{C}\frac{\partial v_2}{\partial t},$$

therefore

$$\frac{K}{C}\frac{\partial^2 i_2}{\partial x^2}-\frac{\partial^2 i_1}{\partial x^2}=\frac{C^2-K^2}{C}\frac{\partial^2 v_1}{\partial x\ \partial t}, \text{ and } \frac{K}{C}\frac{\partial^2 i_1}{\partial x^2}-\frac{\partial^2 i_2}{\partial x^2}=\frac{C^2-K^2}{C}\frac{\partial^2 v_2}{\partial x\,.\,\partial t}. \tag{11}$$

Whence we get from (10) and (11)

$$\frac{K}{C}\frac{\partial^2 i_2}{\partial x^2}-\frac{\partial^2 i_1}{\partial x^2}=\frac{C^2-K^2}{C}\left(L\frac{\partial^2 i_1}{\partial t^2}+M\frac{\partial^2 i_2}{\partial t^2}\right),$$

and

$$\frac{K}{C}\frac{\partial^2 i_1}{\partial x^2}-\frac{\partial^2 i_2}{\partial x^2}=-\frac{C^2-K^2}{C}\left(L\frac{\partial^2 i_2}{\partial t^2}+M\frac{\partial^2 i_1}{\partial t^2}\right)$$

And hence by means of (8)

$$-\frac{M}{L}\left\{\frac{\partial^2 i_2}{\partial x_2}+\frac{K}{M}(L^2-M^2)\frac{\partial^2 i_2}{\partial t^2}\right\}=\left\{\frac{\partial^2 i_1}{\partial x^2}+\frac{K}{M}(L^2-M^2)\frac{\partial^2 i_1}{\partial t^2}\right\}$$

and

$$-\frac{M}{L}\left\{\frac{\partial^2 i_1}{\partial x^2}+\frac{K}{M}(L^2-M^2)\frac{\partial^2 i_1}{\partial t^2}\right\}=\left\{\frac{\partial^2 i_2}{\partial x^2}+\frac{K}{M}(L^2-M)\frac{\partial^2 i_2}{\partial t^2}\right\}. \tag{12}$$

Since it is impossible M^2 should equal L^2, we have therefore,

$$\frac{\partial^2 i_1}{\partial x^2}+\frac{K}{M}(L^2-M^2)\frac{\partial^2 i_1}{\partial t^2}=0 \quad\text{and}\quad \frac{\partial^2 i_2}{\partial x^2}+\frac{K(L^2-M^2)}{M}\frac{\partial^2 i_2}{\partial t^2}=0,$$

therefore

$$\frac{\partial^2 i_1}{\partial x^2}=\frac{C}{L}(L^2-M^2)\frac{\partial^2 i_1}{\partial t^2}\equiv\beta^2\frac{\partial^2 i_1}{\partial t^2}, \text{ and } \frac{\partial^2 i_2}{\partial x^2}=\frac{C}{L}(L^2-M^2)\frac{\partial^2 i_2}{\partial t^2}\equiv\beta^2\frac{\partial^2 i_2}{\partial t^2}. \tag{13}$$

Now equation (13) had the same form as equation (3) and show that each wire, in the presence of the other wire and the screen tube, behaves as if it had a uniform inductance $(L^2-M^2)/L$ per unit length and a capacity C per unit length.

Hence

$$i_1=(A\sin p\beta x+B\cos p\beta x)\cos pt+(D\sin p\beta x+E\cos p\beta x)\sin pt$$

and

$$i_2=(F\sin p\beta x+G\cos p\beta x)\cos pt+(H\sin p\beta x+N\cos p\beta x)\sin pt. \tag{14}$$

These equations are perfectly general and are adaptable to a variety of terminal conditions; for example, the two wires may carry currents from distinct circuits which have no other connecting link. The terminal conditions most interesting to the present discussion are those for which the two wires are connected together at the far end and have one near end connected to the screen tube. Expressed in symbols these conditions are as follows:—

$$x = 0,\quad v_1 = \mathrm{V}\sin pt,\quad i_1 = -\mathrm{I}\cos pt,\quad v_2 = 0,$$

$$x = l,\quad v_1 = v_2,\quad \text{and}\quad i_1 = -i_2,$$

$$i_1 = (\mathrm{A}\sin p\beta x - \mathrm{I}\cos p\beta x)\cos pt$$

$$= -\left(\frac{\mathrm{CV}}{\beta}\sin p\beta x + \mathrm{I}\cos p\beta x\right)\cos pt,\quad \text{since} -\left(\frac{\partial i_1}{\partial x}\right)_{x=0} = \mathrm{C}p\mathrm{V}\cos pt,$$

and

$$i_2 = \left(-\frac{\mathrm{KV}}{\beta}\sin p\beta x + \mathrm{G}\cos p\beta x\right)\cos pt,\quad \text{since} -\left(\frac{\partial i_2}{\partial x}\right)_{x=0} = \mathrm{K}p\mathrm{V}\cos pt,$$

also

$$\frac{\mathrm{CV}}{\beta}\sin p\beta l + \mathrm{I}\cos p\beta l = -\frac{\mathrm{KV}}{\beta}\sin p\beta l + \mathrm{G}\cos p\beta l,\quad \text{since } i_1 = -i_2$$

when $x = l$. Therefore

$$\mathrm{G} = \mathrm{I} + \frac{(\mathrm{C}+\mathrm{K})}{\beta}\mathrm{V}\tan p\beta l,$$

also

$$-\frac{\mathrm{CV}}{\beta}\cos p\beta x + \mathrm{I}\sin p\beta x = -\frac{\mathrm{KV}}{\beta}\cos p\beta x - \mathrm{G}\sin p\beta x,$$

since

$$\left(\frac{\partial i_1}{\partial x}\right)_{x=l} = \left(\frac{\partial i_2}{\partial x}\right)_{x=l},$$

therefore

$$-\mathrm{G} = \mathrm{I} - \frac{(\mathrm{C}-\mathrm{K})}{\beta}\mathrm{V}\cot p\beta l,$$

therefore

$$2\mathrm{I} = \frac{\mathrm{V}}{\beta}\{(\mathrm{C}-\mathrm{K})\cot p\beta l - (\mathrm{C}+\mathrm{K})\tan p\beta l\}.$$

Whence

$$\mathrm{Z} \equiv \frac{\mathrm{V}}{\mathrm{I}} = \sqrt{\frac{\mathrm{L}}{\mathrm{C}}\left(1 - \frac{\mathrm{K}^2}{\mathrm{C}^2}\right)}\,\frac{\tan 2p\beta l}{1 - \mathrm{K}\sec 2p\beta l/\mathrm{C}}, \tag{15}$$

and

$$\mathrm{G} = -\frac{\mathrm{K}/\mathrm{C} - \sec 2p\beta l}{1 - \mathrm{K}\sec 2p\beta l/\mathrm{C}}\,\mathrm{I},$$

therefore,

$$i_1 = -\mathrm{I}\left\{\frac{\cos p\beta l\,(2l - x) - \mathrm{K}\cos p\beta x/\mathrm{C}}{\cos 2p\beta l - \mathrm{K}/\mathrm{C}}\right\}\cos pt. \tag{16}$$

Now

$$\frac{\partial q_1}{\partial t} = -\frac{\partial i_1}{\partial x},$$

therefore

$$q_1 = \frac{\mathrm{I}\beta\,\{\sin p\beta\,(2l - x) + \mathrm{K}\sin p\beta x/\mathrm{C}\}}{\cos 2p\beta l - \mathrm{K}/\mathrm{C}}\sin pt.$$

After very laborious reduction it may be shown that if the previous equation for i_2 is referred to the origin $x = 2l$, then equation (16) results. Consequently a single equation suffices to represent the current flowing up one cylinder and back by the other cylinder.

We see also from (15) that $\mathrm{Z} = 0$, when $2p\beta l = \pi$, 2π, etc., and $\mathrm{Z} = \infty$, when $\cos 2p\beta l = \mathrm{K}/\mathrm{C}$.

When $\mathrm{Z} = 0$, the system is in resonance and there is a current node at $x = l$; it follows that this happens when $\lambda = 4l$, in contrast with $\lambda = 2l$ for the symmetrical system of figs. 3 or 5. Hence one striking result of earthing the point B in fig. 5 is to halve the lowest frequency at which resonance occurs, and a study of equations (15) and (16) shows that the general effect of earthing the point B is approximately the same as earthing the point D together with doubling the length of the line.

We will now consider the force on the system, and refer to fig. 4; a current i_1 at A and a current $-i_2$ at B will produce a repulsion given by

$$\mathrm{F} = 2i_1 i_2\left\{\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'} - \frac{i_2}{i_1}\,\frac{1}{\mathrm{BB}'}\right\}.$$

Similarly charges q_1 and $-q_2$ will produce an attraction given by

$$\mathrm{F} = 2q_1 q_2\left\{\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'} - \frac{q_2}{q_1}\,\frac{1}{\mathrm{BB}'}\right\}.$$

Since we may express i_2 and q_2 as

$$i_2 = -\mathrm{I}\left\{\frac{\cos p\beta x - \dfrac{\mathrm{K}}{\mathrm{C}}\cos p\beta\,(2l - x)}{\cos 2p\beta l - \dfrac{\mathrm{K}}{\mathrm{C}}}\right\}\cos pt,$$

and

$$q_2 = \mathrm{I}\beta\left\{\frac{\sin p\beta x + \dfrac{\mathrm{K}}{\mathrm{C}}\sin p\beta\,(2l - x)}{\cos 2p\beta l - \dfrac{\mathrm{K}}{\mathrm{C}}}\right\}\sin pt.$$

It follows after considerable reduction that we may express F as

$$\mathrm{F} = \left(\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'}\right)\mathrm{I}^2 \frac{\left(1+\frac{\mathrm{K}^2}{\mathrm{C}^2}\right)\cos 2p\beta l - \frac{2\mathrm{K}}{\mathrm{C}}}{\left(\cos 2p\beta l - \frac{\mathrm{K}}{\mathrm{C}}\right)^2}$$

$$- \frac{\mathrm{I}^2}{\mathrm{BB}'} \frac{\cos 2p\beta x + \frac{\mathrm{K}^4}{\mathrm{C}^2}\cos 2p\beta\,(2l-x)\lambda}{\left(\cos 2p\beta l - \frac{\mathrm{K}}{\mathrm{C}}\right)^2} - \frac{2\mathrm{K}}{\mathrm{C}}\cos 2p\beta\,(l-x),$$

therefore

$$\int_0^l \mathrm{F}\,.\,dx = \frac{\mathrm{I}^2}{\left(\cos 2p\beta l - \frac{\mathrm{K}}{\mathrm{C}}\right)^2} \left| \begin{array}{c} l\left\{\left(1+\frac{\mathrm{K}^2}{\mathrm{C}^2}\right)\cos 2p\beta l - \frac{2\mathrm{K}}{\mathrm{C}}\right\}\left(\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'}\right) \\ -\frac{1}{\mathrm{BB}'}\frac{\left(1-\frac{2\mathrm{K}}{\mathrm{C}}-\frac{\mathrm{K}^2}{\mathrm{C}^2}\right)\sin 2p\beta l + \frac{\mathrm{K}^2}{\mathrm{C}^2}\sin 4p\beta l}{2p\beta} \end{array} \right| . \quad (17)$$

By evaluating (17) we may obtain the correcting factor for any assigned value of p. Since it is not intended to use the instrument at frequencies near that which makes $\cos 2p\beta l = \mathrm{K}/\mathrm{C}$, we shall derive by expansion an approximate expression in which $p^3\beta^3l^3$, etc., are ignored. Then

$$\int_0^l \mathrm{F}\,.\,dx = \mathrm{I}^2l\left\{\left(1 \quad \frac{\mathrm{K}}{\mathrm{C}}\right)^2 - 2p^2\beta^2l^2\right\}^{-2} \left| \begin{array}{c} \left\{\left(1-\frac{\mathrm{K}}{\mathrm{C}}\right)^2 - \frac{\mathrm{C}^2+\mathrm{K}^2}{\mathrm{C}^2}2p^2\beta^2l^2\right\}\left(\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'}\right) \\ -\left(1-\frac{\mathrm{K}}{\mathrm{C}}\right)^2\Big/\mathrm{BB}' \end{array} \right| \text{ approx.}$$

$$= \mathrm{I}^2l\left[\left\{1 + 2p^2\beta^2l^2\left(\frac{\mathrm{C}}{\mathrm{C}-\mathrm{K}} - \frac{\mathrm{C}^2+\mathrm{K}^2}{\mathrm{C}^2-\mathrm{K}^2}\right)\right\}\left\{\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'}\right\} - \frac{1}{\mathrm{BB}'}\right| \text{ approx.}$$

$$= \mathrm{I}^2l\left[\left(\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'} - \frac{1}{\mathrm{BB}'}\right) + \left(\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'}\right)\frac{\mathrm{C}^2-2\mathrm{CK}-\mathrm{K}^2}{(\mathrm{C}-\mathrm{K})^2}2p^2\beta^2l^2\right]$$

$$= \mathrm{I}^2l\left(\frac{1}{\mathrm{D}} - \frac{1}{\mathrm{BA}'} - \frac{1}{\mathrm{BB}'}\right)\left[1 + \left(1+\frac{\mathrm{D}}{\mathrm{BA}'}\right)\left(1 - \frac{2\mathrm{K}^2}{(\mathrm{C}-\mathrm{K})^2}\right)\frac{8\pi^2k^2l^2}{\lambda^2}\right] \text{ approx.,} \quad (18)$$

since BA′ is very nearly equal to BB′. Here, as before, k is the constant of the dielectric and λ is the wave-length in three space corresponding to the frequency $n = p/2\pi$.

In a particular instrument to be considered later, and for which various relevant dimensions have been given already, D/BA′ = 1/200 and K/C = − 0·53, inserting these values we find that

$$F = \frac{I^2 l}{D} \times 0{\cdot}99 \left\{1 + \tfrac{1}{3}\pi^2 k^2 \frac{2^2 l^2}{\lambda^2}\right\}. \tag{19}$$

Comparing this expression with (6B) we find the effect of the tube has been to make the correcting factor for longitudinal current distribution 12·5 per cent. greater than for an unscreened system of twice the length. We are contemplating a system of length 15 cm. immersed in a dielectric of constant 2. Then from (19) the force will be 5·4 per cent. greater than for zero frequency when the wave-length is 10 m. and the frequency is 30 million cycles; also the force will not exceed that for zero frequency by as much as 1 per cent. so long as the wave-length is greater than 25 m and the frequency less than 12 million cycles.

(7) *Constructional Details of an Ammeter for Small Currents.*

The force between the wires is $2I^2/D$ dynes per centimetre of length. We have been contemplating that D should be about 4 mm., and then the force will amount to 0·2 dyne per centimetre for a current of 2 A. It is evident from this figure that the elastic control force must be very small in order to obtain a measurable deflection. The system which has been adopted with success may best be described with reference to fig 8, which shows a longitudinal and transverse section of the instrument. One cylinder of the system is fixed rigidly to an ebonite frame which is fixed to the underside of the cover-plate of the thick copper screen tube. The other cylinder stands on two vertical legs of phosphor-bronze suspension strip which is soldered at one end to the cylinder and at the other end to a copper block fixed to the ebonite frame. These two pieces of strip serve to conduct the current to and from the cylinder, and their stiffness supplies the main part of the elastic restoring force. Current is led to the fixed cylinder by a wire passing through a large ebonite disc fixed to one end of the screen tube. Having passed along the fixed cylinder it passes to the bottom of one of the suspension strip legs and thence through the moving cylinder to a point on the end of the screen tube.

The whole screen tube is filled with paraffin oil which serves a three-fold purpose. The moving cylinder is a hollow aluminium tube with sealed ends. The upward thrust of the oil is much greater than the weight of the tube and in this manner the suspension strips are put in tension and therefore have no tendency to buckle. The oil immersion greatly assists in cooling the suspension strips, where necessarily the current density is very high. Also the

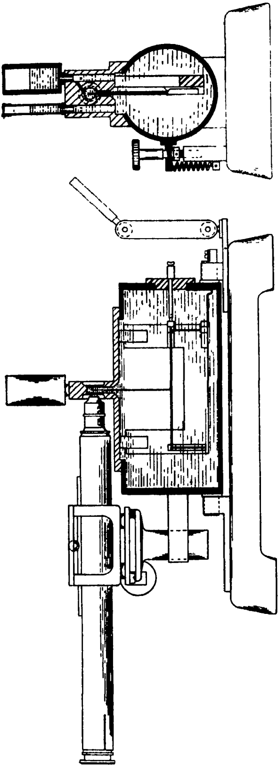


Fig. 8

oil immersion damps the motion of the cylinders and renders robust and portable what would otherwise be a very fragile instrument.

To the middle of the moving cylinder is attached a vertical quartz pointer whose free end passes out through the screen tube, and is viewed through a window in the metal chamber fixed to the cover plate of the screen. The pointer magnifies the motion of the moving cylinder about five times. The movement of the pointer is observed by a microscope with a 1-inch objective and 10-inch tube length, and which has an eye-piece scale 1 cm. long divided into 100 parts. The maximum excursion of the moving cylinder is consequently about 0·20 mm., which is about 5 per cent. of the distance between the centres of the cylinders.

Since there is a considerable upward force of buoyancy on the moving cylinder it is evident the initial distance between the cylinders will depend on the amount by which the plane of the ebonite frame departs from the vertical, and we have here a method of setting the initial distance to a standard value. The whole screen tube is carried on horizontal trunnions so that it can be rotated slightly about an axis parallel to the axis of the tube. This slight rotation is produced by the adjusting screw shown in the end section elevation of fig. 8. To adjust the instrument for use the screen is rotated and the microscope pointer cross traversed simultaneously until the centre line of the pointer and that of a fixed fiducial pointer fixed inside the window is brought to lie on the two terminal divisions of the eye-piece scale; when this adjustment has been effected the cylinders are at their standard distance. Then the microscope is cross traversed until the fixed fiducial pointer has disappeared from view and the quartz pointer has been brought to the zero of the eye-piece scale. The instrument may then be calibrated with steady current.

Since the copper screen tube offers no impediment to the earth's magnetic field, the steady current calibration will generally be affected by the vertical component of this field. For this reason a small permanent magnet is carried on a jointed arm fixed to the case. The position of this magnet is varied until the deflection produced by any given current is independent of its direction, and then the earth's field has been neutralised and the calibration may be made.

Since the frequency correction depends on d^2/D^2, it is important to be assured that D is not less than the value intended, more especially as this distance may get altered unintentionally in the process of assembly. The effect of the earth's vertical field provides a simple method of measuring D after the instrument has been completed. For let the control magnet be removed and a steady current calibration be made with the current passing through the instrument in the direction which gives the smaller deflection for a given current. The current for zero deflection can then be found from this curve; let it be I_0 and the value of the vertical component of field be H. Then $D = 2I_0/H$.

Fig. 9 shows the calibration curve for the instrument described in fig. 8. The eye-piece scale has 100 divisions and the possible accuracy of reading the scale is rather better than 1 per cent. at full deflection; since the scale is sensibly parabolic this corresponds to an accuracy of ½ per cent. of current. It has been shown in section 2 that the equation of the deflection curve is a hyperbola differing little from a parabola. The maximum excursion of the moving cylinder is about 5 per cent. of the initial distance between the cylinders, and from this information it is possible to test whether the curve of fig. 9 fits the hyperbola whose equation is given in section 2. But since the fractional accuracy of reading is poor in the early part of the scale, such

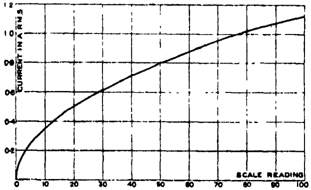


Fig. 9.

a test is hardly justifiable. If for the curve of fig. 9, the deflection from 10 to 100 is plotted against the square of the current, the result is a straight line which does not pass quite through the origin and would make zero current correspond to a positive deflection of one division. Thus, though strictly speaking the scale is hyperbolic, for practical purposes it may be considered parabolic.

Fig. 10 is a calibration curve with steady current and with the earth's magnetic field not neutralised; it illustrates the process of determining I_0.

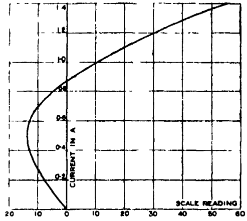


Fig. 10.

Fig. 11 is an external view of the complete instrument.

The heating of the strips and cylinders tends to set up convection currents in the paraffin oil, and these tend to increase the deflection. On switching on the current the deflection attains rapidly a definite value and then tends slowly to creep an extra two divisions at full scale. The current should be deduced from the deflection which results on switching on and before the slow creep has taken place.

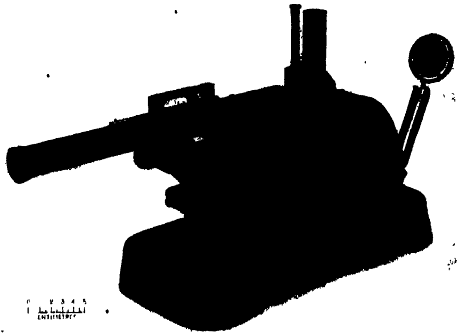


FIG. 11

The instrument described has also a small temperature coefficient, the sensitivity tending to increase slightly with rise of temperature, this may be due in part to the change of density of the paraffin oil. If the instrument is to be used in circumstances where the daily changes of temperature are considerable, it is desirable to calibrate before use. The ammeter here described is probably too sensitive for average requirements, and it is thought that this instrument is probably the most sensitive which can be made in a portable form. In one which would require, say, 5 amperes for full deflection the small inconveniences from convection currents of oil and temperature coefficient probably would disappear. For the control force would then be 25 times as great, while the convection currents and buoyancy forces would remain sensibly the same.

(8) *The Effect of Irregularities at the Ends of the System.*

The arrangement of conductors in fig. 8 departs slightly from the ideal uniform and symmetrically arranged line postulated in fig. 3, mainly because the suspension strip is not in the same straight line with the two cylinders. We must consider the probable effect this will have on the calculated correction factors. The magnetic field will be modified slightly near the sharp corners and this will make a local alteration of current distribution in the cylinders. But since the total correcting factor for proximity effect is only about 3 per cent., an appreciable alteration of this factor over some 5 per cent. of the total length can make only a second order change in the net result.

The effect on the longitudinal current distribution is a little more difficult to predict. That part of the system which is formed by the suspension strip may be regarded as having uniform inductance and capacity per unit length. It is true this portion is perpendicular to the axis of the screen, but since the screen is far away we may reasonably ignore the small effect it has in making the inductance differ from point to point. We may reasonably consider the system as one uniform line, formed by the cylinders, joined to another uniform line formed by the suspension strips; the second line may have slightly different constants per unit length, consequently the effective length of the whole line will be slightly longer or shorter than the actual length; this will make a slight uncertainty in the value of λ/λ_0, but not one sufficient to make appreciable uncertainty in the correcting factor in cases of practical interest when $\lambda > 50\,\lambda_0$.

It follows from equation (3) that if the uniform line is connected in series with a second uniform line of length l' and having constants L′ and C′ per unit length, then

$$i = \mathrm{I}\cos p\alpha\left(l + \frac{\theta}{p\alpha} - x\right)\Big/\cos\,(p\alpha l + \theta),$$

where

$$\tan\theta = p\alpha \mathrm{L}'l'/\mathrm{L}.$$

Since $l' < l$ and $p\alpha l$ is small, θ is small and then

$$i = \mathrm{I}\cos p\alpha\left(l + \frac{\mathrm{L}'}{\mathrm{L}}l' - x\right)\Big/\cos p\alpha\left(l + \frac{\mathrm{L}'}{\mathrm{L}}l'\right)\text{ approx.}$$

The effect of the second uniform line is apparently to increase l in the ratio $(1 + \mathrm{L}'l'/\mathrm{L}l)$. Since L and L′ vary as the logarithm of the distance between the wires it is not possible for L′/L to depart appreciably from unity. Therefore the effective length of the line cannot differ much from the sum of the lengths of the aluminium tube and the suspension strips. Thus there is a very small uncertainty in the ratio of l/λ, but one which will have a second order effect on the correcting factor in cases of practical interest where $\lambda > 50l$. The correcting factor will further be slightly modified because the force between the suspension strips contributes very little to the deflection which arises almost entirely from the force between the aluminium tube and the fixed cylinder.

For this reason we must integrate in formula (23) between $x = 0$ and some value of x which is less than $(l + l')$. In the particular instrument described $l = 10$ cm. and $l' = 6$ cm. If this ratio of l/l' is used, then formula (25) is not altered appreciably except that l inside the brackets must be replaced by $l + l'$. In general the precise value of the correcting factor for longitudinal current distribution is slightly uncertain within very narrow limits, but yet the correction itself is so small even for the shortest wave-length in use at present, that the slight uncertainty is quite unimportant and negligible.

We will now deduce a table of correction factors for the instrument illustrated in fig. 8, and for which the relevant dimensions are as follows: $d = 1 \cdot 1$ mm., $D = 4$ mm., $R = 39$ mm., $l = 10$ cm., $l' = 6$ cm., and $k = 2$.

The correction factor for the screen tube has been found in section 5 and its limiting value is given by formula (8) and has been found to be 1·0 per cent., which has sensibly been attained when the frequency exceeds 125 cycles per second. The correcting factor due to the proximity of the aluminium tube and solid copper cylinder is shown in Table III, its limiting value is 4 per cent., and has sensibly been attained when the frequency exceeds 2 million cycles per second.

It may be calculated that $K/C = -0 \cdot 53$, and hence the correcting factor is infinite when $2p\beta\,(l + l') = 0 \cdot 68\pi$, which is when $\lambda = 12, (l + l')$, if $k = 2$. It will therefore occur for this instrument when $\lambda = 1 \cdot 8$ m.

Table IV.

Frequency n.	5×10^5.	10^6.	10^7.	2×10^7.	3×10^7.
Wave-length in free æther	600 m.	300 m.	30 m.	15 m.	10 m.
Correcting factor for screen tube	0·99	0·99	0·99	0·99	0·99
Correcting factor for cylinders	1·025	1·035	1·04	1·04	1·04
Correcting factor for longitudinal	1·000	1·000	1·006	1·024	1·054
Net factor	1·015	1·025	1·036	1·054	1·084

Fig. 12 shows the results of Table IV plotted as a curve giving the fractional amount by which the observed deflection should be decreased before using it to determine the value of the current from the steady current calibration curve.

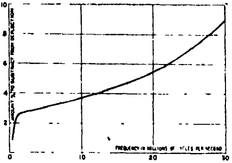


FIG. 12.

(9) *The Resistance and Reactance of the Ammeter.*

We shall define the resistance R of the ammeter as that quantity which when multiplied by the square of the entering current equals the rate of energy dissipation in the system. The resistance will differ from the steady current value because of the distribution of current round the cylinders, and because of the longitudinal current distribution along the cylinders, and because of the eddy currents generated in the screen tube, and because of energy loss in the oil and ebonite dielectric. All except the last item may be calculated.

It may be shown by integrating the expression for σ, given in the note in section (3), that the proximity of the cylinders increases their resistance in the ratio $D/\sqrt{D^2 - d^2}$, when the frequency is very high. This fraction amounts to 1·04 for the specific example. (*Note.*—This expression is well known, see for example Butterworth, 'Phil. Trans.,' vol. 222, p. 72 (1921).) The resistance of an isolated solid circular cylinder is given by the equation

$$R_n/R_0 = \tfrac{1}{4}\sqrt{2}\,(z + 1).$$

For a copper wire 1·1 mm. diameter at 3×10^7 cycles we find that $z = 95$ and so $R_n = 24$, $R_0 = 0 \cdot 5$ ohm per metre. The resistance of the aluminium tube at this frequency will be about 0·7 ohm per metre. Hence the total resistance of the tube and cylinder will be about 0·12 ohm. The energy loss in the shield tube may be derived from the expression for ϕ_0 given in section (5), and may be shown to be quite negligible. The resistance of the very thin suspension strip will not increase appreciably with frequency. The steady current resistance of the complete instrument is 0·2 ohm, and hence it is clear that the resistance of the strips dominates the whole. The effective resistance at 10 m. wave-length will be of the order of 0·3 ohm, and then the power expenditure will be 0·7 W for a current of 1·5 A.

The reactance of the ammeter may be calculated from equation (15) and at 10 m. wave-length will not differ appreciably from the inductive reactance of the system.

It may be shown that

$$2\,(\mathrm{L}-\mathrm{M})\,l = 4l \log \frac{2\mathrm{D}}{d}\left(1-\frac{a^2}{\mathrm{R}^2}\right)$$

$$= 0\cdot128\ \mu\mathrm{H}\ \text{for the instrument considered.}$$

Hence if $n = 3 \times 10^7$ cycles, the reactance of the instrument is 24·5 ohms and then it will require 37 volts for a current of 1·5 amperes.

(10) *An Ammeter for Large Currents.*

It would be an advantage to reduce the inductance of the ammeter. If the left-hand cylinder in fig. 1 was removed and replaced by an infinite conducting plane with its face passing through O and perpendicular to AB, the force on the right-hand cylinder would not be altered but the inductance of the system would be halved. It will now be impossible to calibrate with steady current and we may have to depend entirely on calculated values. Alternatively we may make the diameter of the remaining cylinder so great that the limiting value of the force has sensibly been reached at frequencies which are low enough to permit of calibration by ordinary alternating current instruments. This cannot readily be done for instruments to measure small currents, because the moving parts must be kept light since the operating forces are very small. In these instruments we wish to depend on a steady current calibration, and it is advisable to arrange the distances so that the correcting factor is small and will therefore not be affected appreciably by small irregularities of construction., But if we can construct the ammeter so that it can be calibrated at moderate frequencies, then there is no longer reason to keep the limiting correcting factor small, but merely to arrange that the limiting value is approached closely at low frequencies. The system described as a modification of fig. 1 is unrealisable in practice but may be approached closely by the system illustrated in fig. 13, if R is made very much larger than r. By replacing the system by appropriate image currents it may be shown that the limiting value F of the force per unit length is given by the equation

$$\mathrm{F} = 2d\mathrm{I}^2/\sqrt{(\mathrm{R}^2 + r^2 - d^2)^2 - 4\mathrm{R}^2r^2}.$$

We will make the system definite by assigning values to this expression. Thus, let R = 8 cm., $r = 0{\cdot}8$ cm and $d = 7{\cdot}1$ cm. Then it will be found

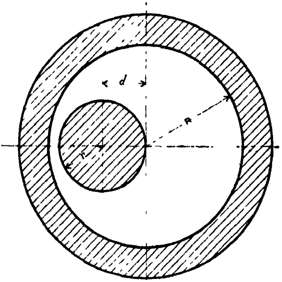


FIG. 13.

that F = 227 dynes for a current of 100 A. If R has been infinite and (R − D) had remained 0·9 cm., it may be found that the force for 100 A would have been 243 dynes, and thus the curvature of the screen reduces the force by only 7 per cent. Thus we may apply formula (2) with confidence to calculate the approximate value of the correcting factor. On doing this we find the force will be within 5 per cent of its limiting value when $z = 7$, which is when $n = 1600$ cycles if the cylinders are of copper. Therefore, by making a calibration curve at various frequencies up to, say, 1000 cycles, it should be possible to infer the correcting factor for all frequencies above a few kilocycles. In the system now proposed the inside cylinder would be connected at one end to the outside return tube, which is also the screen, then in our previous notation $\lambda_0 = 2l$. Hence if $l = 40$ cm., the total force would be about 9000 dynes and the correcting factor for longitudinal current distribution would be 1 per cent. when the frequency was $2{\cdot}5 \times 10^6$ cycles if the cylinder was oil immersed. It may be calculated that the inductance of the system is $0{\cdot}037\ \mu$ H and so would require 58 V to pass 100 A at $2{\cdot}5 \times 10^6$ cycles, and for this current the power absorbed would be of the order of half a watt. The method of supporting the moving cylinder and of observing the deflection would be of the same general character as described already.

The Analysis and Prediction of Tidal Currents from Observations of Times of Slack Water.

By A. T. Doodson, D.Sc., Tidal Institute, University of Liverpool.

(Communicated by J. Proudman, F.R S —Received June 29, 1928.)

1. *Introduction.*

For the purposes of navigation tidal currents may be more important than the rise and fall of the surface, and in fact navigation may be impossible except near times of slack water. In British waters the currents have fairly close connection as regards time and speed with the surface movements, though in some straits in Canadian waters the currents may be largely diurnal in character, whereas the rise and fall of the surface may be semidiurnal in general character. Therefore, the currents cannot be satisfactorily inferred from the tables of high and low water times and heights, and this is probably true in general.

The measurement of tidal currents is not easily effected, but if hourly values are available then the ordinary methods of tidal analysis can be used, though such is rarely the case. The usual method of observation is to determine the time of slack water if the currents are constant in direction.

In this paper the problem of obtaining harmonic constants for the principal tidal constituents representing the current flow, from a knowledge of the times of slack water only, has been solved. The solution is not one of great exactness, but resulting predictions have been considered sufficiently accurate to be included in standard tide tables. The times of maximum current in either direction can also be predicted from the same constants, but velocities of maximum current can only be stated on an arbitrary scale unless a few values of maximum currents have been observed.

It may be noted that this method can similarly be used to obtain fairly good approximations to the tidal elevation at a place, from the times of high and low water, or from the times of half-tide—the latter, for preference, as they can be accurately obtained.

The actual methods of computation are similar to those used in the author's method of analysis of hourly tidal observations. This method has been published in the Transactions of the Royal Society* ; similar methods have been devised for use with short lengths of record, and these have been published by H.M. Stationery Office, for the Hydrographic Department of the Admiralty, for official use. It has not been considered necessary, therefore, to enter into detailed explanation or justification of the numerical methods of analysis used in this paper, and frequent reference is made to the two papers just mentioned.

---

* 'Phil. Trans.,' A, vol. 227, p. 223 (1928).

Two methods are given, one for observations of all slack waters, and one for daylight observations only. The former is illustrated by an example.

2. *General Principles.*

It will be supposed that the current velocity u is capable of representation by a series of harmonic constituents typified by $\mathrm{R}\cos(\sigma t-\epsilon)$, where R is the amplitude, σ is the speed, and $-\epsilon$ is the phase at the origin of time. It will also be supposed that there is a predominant constituent, such as M_2, the degree of predominance remaining undefined at present; a zero suffix will be used to indicate this constituent.

Then if $\sum\limits_c$ is a symbol denoting the aggregation of all constituents other than the predominant one, we have

$$\begin{aligned} u &= \mathrm{R}_0\cos(\sigma_0 t-\epsilon_0)+\sum_c \mathrm{R}\cos(\sigma t-\epsilon) \\ &= (\mathrm{R}_0+\sum_c \mathrm{R}\cos\phi)\cos(\sigma_0 t-\epsilon_0)-\sum_c \mathrm{R}\sin\phi\sin(\sigma_0 t-\epsilon_0), \end{aligned}$$

where

$$\phi=(\sigma-\sigma_0)\,t-(\epsilon-\epsilon_0).$$

Slack water occurs when

$$\sigma_0 t=\epsilon_0+m\pi+\tfrac{1}{2}\pi+\eta,$$

where m is a positive integer and

$$\tan\eta=-\sum_c \mathrm{R}\sin\phi/(\mathrm{R}_0+\sum_c \mathrm{R}\cos\phi).$$

This can be expanded in a convergent series provided that $\mathrm{R}_0>\sum\limits_c \mathrm{R}\cos\phi$, which gives one criterion for the degree of predominance. Writing

$$\sum_c \mathrm{R}\cos\phi=\mathrm{R}_0\gamma\cos\theta \quad\text{and}\quad \sum_c \mathrm{R}\sin\phi=\mathrm{R}_0\gamma\sin\theta,$$

temporarily, then it is readily shown that

$$\eta=-\gamma\sin\theta+\tfrac{1}{2}\gamma^2\sin 2\theta-\tfrac{1}{3}\gamma^3\sin 3\theta+\dots.$$

Taking only the first two terms of this series we obtain

$$\eta = -\sum_{c} \frac{\mathrm{R}}{\mathrm{R}_0} \sin \phi + \sum_{c} \frac{\mathrm{R}}{\mathrm{R}_0} \sin \phi \sum_{c} \frac{\mathrm{R}}{\mathrm{R}_0} \cos \phi$$
$$= -\sum_{p} \frac{\mathrm{R}_p}{\mathrm{R}_0} \sin \phi_p + \left[\sum_{s} \tfrac{1}{2} \frac{\mathrm{R}_s{}^2}{\mathrm{R}_0{}^2} \sin 2\phi_s + \sum_{r,s} \frac{\mathrm{R}_r\mathrm{R}_s}{\mathrm{R}_0{}^2} \sin (\phi_r \pm \phi_s)\right],$$

where we have introduced the suffix p to denote "primary terms" and suffixes r, s to denote "secondary terms"—the former occur in the expression for u while some of the latter will not.

Denoting the time-rate of increase of ϕ by $\dot{\phi}$ we remark that we can ignore the secondary terms in our subsequent analysis unless $2\phi_s$ or $\phi_s \pm \phi_r$ happen to coincide in any instance with a value of ϕ_p.

A table of values of μ, which is the phase-increment in degrees per lunar day, and which is strictly proportional to $\sigma - \sigma_0$, is given in Table I, and it is clear

Table I.—Values of σ and μ for the Most Important Tidal Constituents.

Const.	σ.	μ.	Const	σ.	μ.	Const.	σ.	μ.
	°	°		°	°		°	°
Q_1	13 39866	−27·16116	μ_2	−27·96821	−25·23607	MO_3	42·92714	−13 63823
O_1	13 94304	−13 63823	N_2	28·43973	−13 52292	M_3	43 47616	0·00000
P_1	14·95893	11·59784	ν_2	28·51258	−11 71316	MK_3	44·02517	13·63823
S_1	15 00000	12·61804	M_2	28·98410	0 00000			
K_1	15·04107	13·63823	λ_2	29·45563	11 71316	M_4	57·96821	0·00000
J_1	15·58544	27 16116	L_2	29·52848	13·52292	MS_4	58·98410	25·23607
			T_2	29·95893	24 21588			
			S_2	30·00000	25 23607			
			K_2	30·08214	27 27646			

σ is given in degrees per mean solar hour and μ is given in degrees per mean lunar day, multiples of 360° being omitted.

that we have no appreciable perturbations arising from terms with arguments $2\phi_s$. The following table gives some idea of the perturbations that exist :—

Primary constituent (p)	M_2	M_2	M_2	M_2	S_2	S_2
Disturbing constituents (r, s)	N_2, L_2	ν_2, λ_2	S_2, μ_2	K_1, O_1	N_2, λ_2	ν_2, L_2
Relative magnitudes ($\mathrm{R}_r\mathrm{R}_s/\mathrm{R}_0\mathrm{R}_p$)	0·01	0 00	0 01	0 08	0 00	0·00

Primary constituent (p)	S_2	N_2	ν_2	L_2	K_1
Disturbing constituents (r, s)	P_1, K_1	λ_2, S_2	L_2, S_2	ν_2, S_2	S_2, P_1
Relative magnitudes ($\mathrm{R}_r\mathrm{R}_s)/\mathrm{R}_0\mathrm{R}_p$)	0 08	0·02	0·47	0 47	0·16

The relative magnitudes are based upon the relative equilibrium values within each species, with $\mathrm{K}_1 = \frac{1}{2}\mathrm{M}_2$. No important perturbations of J_1 and Q_1 appear to exist.

It is now clear that even if the diurnal tide is fairly large relatively to the semidiurnal tide we shall obtain quite good approximations to the harmonic constants for the principal constituents by ignoring the secondary terms. Having shown this possibility, we now proceed to develop the method of procedure.

3. *Harmonic Expansion for Slack-Water Times.*

Discarding the use of suffixes r, s, p of the preceding paragraph, and writing

$$\xi = \mathrm{R}/\mathrm{R}_0 \tag{3.1}$$

then, to the degree of approximation just indicated,

$$\sigma_0 t = \varepsilon_0 + m\pi + \tfrac{1}{2}\pi + \sum_{\sigma} \xi \cos\{(\sigma - \sigma_0)\,t - (\varepsilon - \varepsilon_0) + \tfrac{1}{2}\pi\}, \tag{3.2}$$

and with exactly the same character of approximation as before, we can write

$$\sigma_0 t = \varepsilon_0 + m\pi + \tfrac{1}{2}\pi$$

on the right of (3.2). We also find it convenient to consider four separate sequences of slack waters defined by $s = 0, 1, 2, 3$, with

$$m = s + 4\mathrm{L},$$

where L obviously denotes the number of complete lunar days elapsed since the origin of time; thus in each sequence we deal with slack waters at intervals of approximately a lunar day.

Now let p denote the species number of the constituent, as usual, so that σ is approximately $\tfrac{1}{2}p\sigma_0$. Since the number of hours in a lunar day is $4\pi/\sigma_0$ then the phase-increment in a lunar day is

$$\sigma \times 4\pi/\sigma_0 = 2p\pi + \mu,$$

where μ is defined as in § 2, multiples of 2π being ignored. Then we can write

$$\pi s \frac{\sigma - \sigma_0}{\sigma_0} = \tfrac{1}{4}s\mu + s\,(p-2)\,\tfrac{1}{2}\pi = \Delta_s, \tag{3.3}$$

and writing also

$$\delta = \varepsilon - \sigma\,(\varepsilon_0 + \tfrac{1}{2}\pi)/\sigma_0 \tag{3.4}$$

we obtain finally

$$\sigma_0 t - 4\pi\mathrm{L} = \varepsilon_0 + s\pi + \tfrac{1}{2}\pi + \sum_{\sigma} \xi_s \cos\{\mu\mathrm{L} + \Delta_s - \delta_s\} = \sigma_0\mathrm{D}_s \tag{3.5}$$

so defining D_s. The suffixes to ξs and δs are attached only for the purpose of numerical convenience.

If we define our origin of time as the time of a particular lower transit of the "mean moon" at Greenwich, then successive lower transits occur at intervals of $4\pi/\sigma_0$ mean solar hours, and therefore D_s is the interval of time between a lower transit of the mean moon and slack water. It is immaterial to our purpose whether the transit is taken at Greenwich or at the place of observation, or whether the time is G.M.T. or Standard Time. It is very convenient, however, to deal only with transits of the mean moon at Greenwich, to measure them in G.M.T. and to express the observations in Standard Time appropriate to the place. The difference D_s then becomes the difference between two numbers of hours, but it is convenient to refer to the values of D_s as being expressed in hours—in fact, we shall take the observations as they stand and treat them as though they were in G.M.T. and had been taken at Greenwich.

There is no special definition of the sequences except that the sequence for $s = 3$ is about six hours later than the sequence $s = 2$, and this again is about six hours later than the sequence $s = 1$, which also is about six hours later than sequence 0. To avoid negative quantities in D_0 it is convenient to take the sequence $s = 0$ so that its mean value lies between 0 and 6.

4 *A Constant Current.*

Since a constant current is frequently of great importance provision must be made for it in the analysis. Putting $\sigma = \mu = \epsilon = p = 0$ and $\xi = C$, then $\Delta_s = -s\pi$ and the contribution of the constant current to D_s is $(-1)^s\, C/\sigma_0$.

5. *Tidal Constituents* M_1, M_3, M_4.

These yield contributions to $\sigma_0 D_s$ defined by

$$\xi_s \cos(-\tfrac{1}{2}\pi s + \Delta_s - \delta_s), \quad \xi_s \cos(\tfrac{1}{2}\pi s + \Delta_s - \delta_s), \quad \xi_s \cos(\pi s + \Delta_s - \epsilon_s)$$

respectively. Obviously M_1 and M_3 will not be separable one from the other. These constituents can be left out of account for the purposes of analysis and prediction. M_4 will not be distinguishable analytically from a constant current.

6. *Conjugate Constituents.*

There are certain pairs of constituents which are conjugate to one another in the sense that they have equal and opposite values of μ.

Such conjugates are K_1 and O_1; let their values of δ and ξ be denoted by single and double dashes respectively, and suppose that the analysis is carried out as for K_1. Then, if O_1 is entirely absent we should get (theoretically) the same value for δ' from each of the four sequences $s = 0, 1, 2, 3$. But if K_1 is entirely absent the analysis (as for K_1) will yield $-\delta''$ from sequences 0, 2, and

$180° - \delta''$ from sequences 1, 3. Hence, if both K_1 and O_1 are present, we have

$$\begin{array}{lll} (s = 0, 2) & \text{apparent } \xi\cos\delta & = (\xi\cos\delta)' + (\xi\cos\delta)'' \\ (s = 1, 3) & \quad ,, \quad \xi\cos\delta & = (\xi\cos\delta)' - (\xi\cos\delta)'' \\ (s = 0, 2) & \quad ,, \quad \xi\sin\delta & = (\xi\sin\delta)' - (\xi\sin\delta)'' \\ (s = 1, 3) & \quad ,, \quad \xi\sin\delta & = (\xi\sin\delta)' + (\xi\sin\delta)''. \end{array}$$

Thus it is possible to separate the conjugates K_1, O_1, and, similarly, the conjugates J_1, Q_1 can be separated. Conjugates of the semidiurnal species ($p = 2$) cannot be separated from sequences of slack waters; in particular, the analysis for N_2 will yield results containing contributions from L_2, and that for S_2 will yield results perturbed by μ_2. The inability to separate conjugates is of no moment if predictions of slack-water times only are required, but if variations of current strength are desired then it will be necessary to make some assumption as to the relative contributions of the conjugates, just as an allowance for K_2 is made when analysing for S_2 from a short length of record.

7. *Correlation with Tide-Generating Potential.*

If V is the phase of the constituent of the tide-generating potential at Greenwich at the origin of time, f and u are the usual nodal corrections, also taken at the origin of time (which is taken as the central day of the observations treated), then Harmonic Constants (H and g) are obtained from the formulæ

$$H = R/f, \qquad g = V + u + \epsilon.$$

Since $\xi = R/R_0$ then, if we take $R_0 = f_0$ so that $H = 1$ for the predominant constituent M_2, we obtain

$$H = \xi f_0/f.$$

If the observations have been taken in Standard Time S hours slow on Greenwich, and if the Place is in longitude L° west of Greenwich, we have

$$g = \kappa + pL - \sigma S,$$

where κ is the usual form of the phase-lag, according to the usage of Darwin. In practice there is nothing to be gained by evaluating κ, but in all cases g should not be quoted without the values of S and L, and, above all, it should never be referred to as "kappa" or "kappa in standard time."

It should be noted that it is necessary to specify the direction of the current in some way; *e.g.*, "current turning from north to south"; see Table VIII.

8. *Times of Lower Transits of the Mean Moon.*

If h and s are the mean longitudes in degrees, of the sun and moon, respectively, at mean solar midnight at Greenwich, then the mean solar time, in hours, of the lower transit of the "mean moon" is given by

$$(s - h + 180°) \div 14°{\cdot}492, \tag{8.1}$$

and successive transits occur at intervals of 24·8412 mean solar hours.

9. *Methods of Computation.*

The analysis is best carried out in "months" of 29 lunar days, and we shall denote the central day of a month by $\bar{L}$, so that $L - \bar{L}$ varies from -14 to 14. Writing

$$\sigma_0 D_s = \epsilon_0 + s\pi + \tfrac{1}{2}\pi + \Sigma\, \xi_s \cos\{\mu(L - \bar{L}) + \mu\bar{L} + \Delta_s - \delta_s\}, \tag{9.1}$$

then our object is to combine the 29 values of D_s in such a way that particular constituents are magnified relatively to others. To do this, we place alongside the 29 values of D_s a strip of paper containing 29 multipliers d_1, as given in Table II, and the sum of the products, term by term, is called D_{s1}. The second suffix indicates the multipliers which have been used. Since the multipliers d_0, d_1, d_2 are integers approximately proportional to $2\cos\mu(L - \bar{L})$ with μ approximately equal to 0°, $-12°$, $-24°$ respectively, it is readily shown that D_{s1} contains very small contributions from constituents whose values of μ are not approximately equal to $\pm 12°$, while D_{s2} magnifies, relatively to the rest, constituents whose values of μ are approximately $\pm 24°$. Similarly d_a, d_b are integers approximately representing $2\sin\mu(L - \bar{L})$ with μ approximately equal to $-12°$ and $-24°$ respectively.

Table II.—List of Daily Multipliers d_0, d_1, d_a, d_2, d_b.

$L - \bar{L}$	−14	−13	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
d_0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
d_1	−2	−2	2	2	2	1	1	1	1	0	0	−1	−1	−1	−2	−2	−2
d_a	0	0	0	−1	−1	−1	−1	−2	−2	−2	−2	−2	−2	−1	−1	−1	0
d_2	2	2	2	2	1	1	0	−1	−2	−2	−2	−2	−1	0	1	2	2
d_b	−1	−1	0	−1	−1	−2	−2	−2	−1	−1	1	1	2	2	2	1	1
Sum	0	1	5	3	2	0	−1	−3	−3	−4	−2	−3	−1	1	1	1	2

d_1, d_2 are symmetrical about $L - \bar{L} = 0$. d_a, d_b are asymmetrical about $L - \bar{L} = 0$.

Since the sequence $s = 2$ is approximately 12 hours later than the sequence $s = 0$ it follows that semidiurnal contributions to D_0 and D_2 have approximately equal phases, while diurnal contributions differ in phase by 180° in the two sequences. Hence $D_0 + D_2$, $D_0 - D_2$ will contain very small contributions from diurnal and semidiurnal constituents respectively, and we make use of these combinations in order to separate the constituents. In actual practice we find it most convenient to compute D_{s0}, D_{s1}, D_{sa} … for each sequence and then to form such combinations as

$$D_{02} + D_{22}, \quad D_{02} - D_{22}.$$

From (9.1) it is easily shown that

$$\sigma_0 D_{s2} = \Sigma\, \xi_s a_2 \cos(\delta_s - \Delta_s - \mu\bar{L})$$

where a_2 is a factor depending entirely on the effects of the multipliers d_2 upon $\cos \mu(L - \bar{L})$; that is,

$$a_2 = \Sigma\, d_2 \cos \mu(L - \bar{L}) \quad (L - \bar{L} = -14 \text{ to } 14),$$

and clearly it is possible to determine a_2 once for all, it is only necessary in the present method to evaluate it for the principal contributing functions, and using exact values of $\cos \mu(L - \bar{L})$ the values of a_2 have been found, together with analogous factors associated with the use of d_1, d_a …. We note that δ_s and ξ_s for a particular constituent should be constants for all sequences, and we have used the suffix simply to indicate from what sequence they have been obtained. Now

$$\cos(\delta_0 - \Delta_0 - \mu\bar{L}) + \cos(\delta_2 - \Delta_2 - \mu\bar{L}) = 2\cos\Delta_1 \cos(\delta_1 - \Delta_1 - \mu\bar{L}),$$

since $\tfrac{1}{2}(\Delta_0 + \Delta_2) = \Delta_1$, and therefore

$$\sigma_0(D_{s0} + D_{s2}) = \Sigma\, 2\cos\Delta_1\, a_2 \xi_2 \cos(\delta_1 - \Delta_1 - \mu\bar{L}).$$

Such factors as $2a_2 \cos \Delta_1 \div \sigma_0$ have been computed for the principal constituents and we have obtained the following results:

$$\left.\begin{array}{lll}
D_{02} + D_{22} = & 123{\cdot}27\, \xi_1 \cos(\delta_1 - \Delta_1 - \mu\bar{L}) & \multirow{2}{*}{$\Big\}\; S_2 : \mu = 25^\circ{\cdot}236$} \\
D_{0b} + D_{2b} = & -113{\cdot}82\, \xi_1 \sin(\delta_1 - \Delta_1 - \mu\bar{L}) & \\
D_{01} + D_{21} = & 114{\cdot}79\, \xi_1 \cos(\delta_1 - \Delta_1 - \mu\bar{L}) & \Big\}\; N_2 : \mu = -13^\circ{\cdot}523 \\
D_{0a} + D_{2a} = & 109{\cdot}47\, \xi_1 \sin(\delta_1 - \Delta_1 - \mu\bar{L}) & \\
D_{0a} - D_{2a} = & 109{\cdot}47\, \xi_1 \cos(\delta_1 - \Delta_1 - \mu\bar{L}) & \Big\}\; K_1 : \mu = 13^\circ{\cdot}638 \\
D_{01} - D_{21} = & 114{\cdot}79\, \xi_1 \sin(\delta_1 - \Delta_1 - \mu\bar{L}) & \\
D_{0b} - D_{2b} = & 98{\cdot}84\, \xi_1 \cos(\delta_1 - \Delta_1 - \mu\bar{L}) & \Big\}\; J_1 : \mu = 27^\circ{\cdot}161 \\
D_{02} - D_{22} = & 118{\cdot}11\, \xi_1 \sin(\delta_1 - \Delta_1 - \mu\bar{L}) & .
\end{array}\right\} \quad (9.2)$$

The suffixes attached to ξ_1, δ_1, Δ_1 indicate that by combining D_2 and D_0 we have averaged the values from $s = 2$ and $s = 0$. Similar formulæ arise from D_2 and D_1 with ξ_2, δ_2 and Δ_2 on the right.

It is necessary to remember that the apparent constants for S_2 will have to be corrected for K_2, that ν_2 perturbs N_2, and that both ρ_1 and O_1 contribute to the same functions as K_1.

Allowing for the constant current we have

$$\sigma_0 D_{s0} = 29\,(\epsilon_0 + s\pi + \tfrac{1}{2}\pi) + 29C\,(-1)^s/R_0.$$

It is convenient to express ϵ_0 in degrees and therefore we obtain

$$D_{s0} = (\epsilon^\circ + s180^\circ + 90^\circ) + 57\cdot3\ C\,(-1)^s/R_0. \qquad (9.3)$$

The results of analysis for each month should be combined vectorially in the usual manner; that is, average values of H cos g, H sin g should be obtained.

10. *Examples of Calculations.*

We shall illustrate the methods of computation by means of an example for Turn Point, Boundary Pass, B.C., Canada. Summer observations of slack-water times for the year 1926 were sent for analysis, and these times, in hours and minutes of standard time 8 hours slow on Greenwich, were reduced to hours and decimals. Table III gives a specimen table, which also shows the values of the G.M.T. of lower transit of the mean noon, deduced from formula (8.1). For values of s and h reference should be made to the author's paper already cited or to tables in common use.

Table III.—Observations and Calculations of D. Turn Point.

Date.	G.M.T. of lower transit of mean moon.	" H.W. slack."	D.	" H.W. slack."	D.	" L.W. slack."	D.	" L.W. slack."	D.
July 17	5·63	10·00	4·37	23·20	17·57	5·15	—	16·13	10·50
,, 18	6·47	11·11	4·64	23·85	17·38	6·11	24 48	16·76	10·29
,, 19	7·31	—	—	12·91	5·60	6·95	24·48	17·83	10·52
,, 20	8·15	0·36	17·05	14·20	6·05	7·71	24·40	19·00	10·85
Aug. 7	23·31	3·56	5·09	17 85	19·38	10·91	12·44	23·91	25·44
,, 8	—	4·15	4·84	18·36	19·05	11·65	12·34	—	—
,, 9	0·14	4·50	4·36	19 00	18·86	0·45	25·14	12·20	12·06

The observations of slack waters were called " H.W. slack " when they were closely related in time to high water at a neighbouring place.

The computation of D requires care and the example has been carefully chosen to illustrate all the pitfalls for the unwary; the trouble arises from the mixture of solar and lunar time; for example, on July 20, H.W. slack is at 0·36—this time should be considered as 24·36, and 7·31 (July 19) should be subtracted to give D.

Table IV gives values of D_0, D_1, D_2, D_3 for a month, with $L = 0$ on July 31, 1926. The value of $\bar{L}$ is taken as zero. In practice the 29 values of each function are arranged in a single column.

Table IV.—Values of D_0, D_1, D_2, D_3. Turn Point, $\bar{L} = 0$ on July 31, 1926.

$L - \bar{L}$.	D_0	D_1	D_2	D_3.	$L - L$.	D_0.	D_1.	D_2.	D_3.
−14	4·37	10 50	17 57	24 48	1	5 39	12 84	19 12	23·75
−13	4·64	10 29	17·38	24·48	2	5·15	13 11	19 93	24·73
12	5·60	10 52	17 05	24·40	3	5·16	13 11	19·99	25·41
−11	6 05	10·85	17 03	24·63	4	5 08	12·80	19·92	25 65
−10	6·41	11·26	16 85	24 45	5	5·10	12 67	19·55	25·60
9	6·24	11·79	16·66	24 21	6	5 00	12 44	19·38	25·44
−8	6 20	11·97	16 55	24 37	7	4·84	12·34	19·05	25·14
−7	6 14	12·08	16 04	24·31	8	4 36	12 06	18·86	24·92
6	5·85	11·87	15 99	24 10	9	4·41	11·64	18·84	24·76
−5	5·83	12·10	16 06	23 83	10	4·18	11 48	18·55	24·52
−4	6·22	12·01	16 41	23 81	11	3 93	11 21	18 19	24·31
3	5·78	12·18	16 62	23·78	12	[4 00]	11 02	17·77	24 04
−2	5·42	12·22	17 03	23 38	13	4·04	10 74	17·47	23 80
−1	5·69	12 24	17 51	23 38	14	4 56	10 50	16 95	23 88
0	5 69	12 65	18 30	23 57					

Note.—The square bracket indicates an interpolation

Table V gives the values of D_{s0}, D_{s1}, D_{sa}, D_{s2}, D_{sb}.

From the four values of D we deduce according to formula (9.2)

$$\epsilon_0 = 69^\circ\cdot 7, \qquad C = -0\cdot 13\, R_0. \tag{10.1}$$

Table V.—Values of D_{s0}, D_{s1}, D_{sa}, D_{s2}, D_{sb}. Turn Point, $\bar{L} = 0$ on July 31, 1926.

D_{00}	151·4	D_{10}	342·5	D_{20}	516 6	D_{30}	707·2
D_{01}	13 2	D_{11}	29·9	D_{21}	13 2	D_{31}	− 1·8
D_{0a}	24 0	D_{1a}	−7·6	D_{2a}	−45·5	D_{3a}	−14·1
D_{02}	−7·9	D_{12}	−8·2	D_{22}	3·2	D_{32}	−12 5
D_{0b}	−8·4	D_{1b}	−4 4	D_{2b}	−20·5	D_{3b}	−15·8

Table VI gives values of $\xi_1 \cos(\delta_1 - \Delta_1)$, etc., as deduced from (9.2). From these it is a straightforward matter to complete the calculations of that table. The values of Δ_1 and Δ_2 are the same for all analyses, though not for all constituents. The calculation of $\xi_1 \cos \delta_1$, etc., is necessary for the separation of K_1 and O_1, J_1 and Q_1.

Table VI.—Computation of ξ and δ.—Turn Point, $\overline{L} = 0$ on July 31, 1926.

Principal constituent	K_1.	J_1	S_2.	N_2.
$\xi_1 \cos(\delta_1 - \Delta_1)$	0·635	0·122	−0·035	0·230
$\xi_1 \sin(\delta_1 - \Delta_1)$	0·000	−0·094	0·254	−0·196
ξ_1	0·635	0·154	0·256	0·303
$\delta_1 - \Delta_1$	0·0	322·4	97·8	319·6
Δ_1	273·4	276·8	6·3	− 3·4
δ_1	273·4	239·2	104·1	316·2
$\xi_2 \cos(\delta_2 - \Delta_2)$	0·059	0·115	−0·168	0·244
$\xi_2 \sin(\delta_2 - \Delta_2)$	0·276	0·036	0·177	−0·198
ξ_2	0·284	0·121	0·245	0·314
$\delta_2 - \Delta_2$	77·9°	17·4°	133·5°	321·0°
Δ_2	186·8	193·6	12·6	− 6·8
δ_2	264·7	211·0	146·1	314·2
$a = \xi_1 \cos \delta_1$	0·038	−0·079	−0·062	0·219
$b = \xi_2 \cos \delta_2$	−0·026	−0·104	−0·204	0·219
$c = \xi_1 \sin \delta_1$	−0·633	−0·132	0·249	−0·210
$d = \xi_2 \sin \delta_2$	−0·282	−0·062	0·136	−0·225
$\frac{1}{2}(a + b) = \xi \cos \delta$	0·006	−0·092	−0·133	0·219
$\frac{1}{2}(c + d) = \xi \sin \delta$	−0·458	−0·097	0·193	−0·218
ξ	0·46	0·13	0·23	0·31
δ	270·8°	226·5°	124·6°	315·2°
Conjugate constituent.	**O_1.**	**Q_1.**		
$\frac{1}{2}(a - b) = \xi \cos \delta$	0·032	0·013		
$\frac{1}{2}(d - c) = \xi \sin \delta$	0·176	0·035		
ξ	0·18	0·04		
δ	79·7°	69·6°		

To correct K_1 for P_1 use must be made of the calculations of $V + u$ in Table VII. Let f be the value of f for K_1, and let

$$W = (V \text{ of } P_1) - (V + u \text{ of } K_1).$$

Then we define

$$\left.\begin{aligned} wf &= \tan^{-1}\{0{\cdot}33 \sin W/(1 + 0{\cdot}33 \cos W)\} \\ 1 + rf &= (1{\cdot}11 + 0{\cdot}66 \cos W)^{\frac{1}{2}} \end{aligned}\right\}, \qquad (10.2)$$

whence we deduce w, and r.

Table VII.—Corrections for K_1, S_2, N_2. Turn Point, $\overline{L} = 0$ on July 31, 1926. ($V + u$ is computed in Table VIII.)

K_1.		S_2.		N_2.	
f of K_1	$= 0 \cdot 98$	f of K_2	$= 0 \cdot 94$		
$V + u$ of P_1	$= 44 \cdot 2^\circ$	$V + u$ of K_2	$= 23\ 5^\circ$	$V + u$ of ν_2	$= 249 \cdot 6^\circ$
$V + u$ of K_1	$= 111 \cdot 1$	$V + u$ of S_2	$= 162 \cdot 6$	$V + u$ of N_2	$= 304 \cdot 0$
W	$= 293 \cdot 1$	W	$= 220\ 9$	W	$= 305 \cdot 6$
wf	$= -15 \cdot 1$	w/f	$= -12 \cdot 6$		
w	$= -14 \cdot 8$	w	$= -11 \cdot 9$	w	$= -7 \cdot 8$
rf	$= 0 \cdot 17$	r/f	$= -0\ 19$		
$1 + r$	$= 1\ 17$	$1 + r$	$= 0\ 80$	$1 + r$	$= 1 \cdot 12$

To correct S_2 we take f as the value of f for K_2 and let

$$W = (V + u \text{ of } K_2) - (V \text{ of } S_2).$$

Then we define

$$\left.\begin{aligned} w/f &= \tan^{-1}\{0 \cdot 27 \sin W/(1 + 0 \cdot 27 \cos W)\} \\ 1 + r/f &= (1 \cdot 07 + 0 \cdot 54 \cos W)^{\frac{1}{2}} \end{aligned}\right\}. \qquad (10.3)$$

To correct N_2 we take

$$W = (V + u \text{ of } \nu_2) - (V + u \text{ of } N_2)$$

and then define

$$\left.\begin{aligned} w &= \tan^{-1}\{0 \cdot 19 \sin W/(1 + 0 \cdot 19 \cos W)\} \\ 1 + r &= (1 \cdot 04 + 0 \cdot 38 \cos W)^{\frac{1}{2}} \end{aligned}\right\}. \qquad (10.4)$$

In all cases w must be added to δ, and R must be divided by $1 + r$. An example is given in Table VIII.

The final computations for H and g are of a well-known nature. The exact origin of time for which V, u, f are computed is the G.M.T. of the lower transit of the mean moon at Greenwich on the central day ($L = 0$). The computations for M_2, S_2, K_1, O_1 are set out below. For M_2, to save special directions, we write $\delta = 270^\circ$ always. The value of H for M_2 is taken to be unity so that ξ for $M_2 = f_0$.

The value of g_0 indicates that the maximum of the curve occurs about 2½ hours after the transit, and therefore the first slack water after the maximum is about 5½ hours after transit; hence it must in this instance correspond to sequence $s = 0$ or 2.

Table VIII.—Computation of H and g. Turn Point, S = 8.

Origin of Time at 17·42 hours on July 31, 1926, G.M.T.

$\varepsilon_0 = 69{\cdot}7^\circ, \quad C = -0{\cdot}13 f_0 = -0{\cdot}13.$

V, u, f are computed for the origin of time.

	M_2.	S_2.	K_1.	O_1.	
δ	270·0°	124·6°	270·8°	79·7	
$(\epsilon_0 + 90^\circ)\,\sigma/\sigma_0$	159·7	165·2	82·6	76·7	
V	0·0	162·6	119·9	240·2	
u	− 2·1	0·0	− 8·8	11·0	
w		− 11·9	− 14·8		
sum = g	67·6	80·5	89·7	47·6	
ξ	1·01	0·23	0·46	0·18	
$1 + r$		0·80	1·17		
f	1·01	1·00	0·98	0·98	
$f_0\xi + f(1 + r) = H$	1·00	0·29	0·40	0·18	

The first slack water after maximum of the curve obtained with these constants corresponds to one of the sequences of slack waters denoted by " H.W. slack " in the observations.

11. *Second Approximations.*

When first approximations to the values of the constants have been obtained the predicting machine may be used to compare " predictions " with observations and the differences (O − P) may be analysed. Let single dashes refer to the results of the first analysis and let double dashes refer to the results of the second analysis. Then we can write

$$D_s = D_s' + D_s'', \tag{11.1}$$

and since

$$\sigma_0 D_s = \varepsilon_0 + s\pi + \tfrac{1}{2}\pi + \sum_c \xi_w \cos(\mu L + \Delta_s - \delta_s),$$

$$\sigma_0 D_s' = \varepsilon_0' + s\pi + \tfrac{1}{2}\pi + \sum_c \xi_w' \cos(\mu L + \Delta_s - \delta_s'),$$

then

$$\sigma_0 D_s'' = \varepsilon_0'' + \sum_c \xi_w'' \cos(\mu L + \Delta_s - \delta_s'') \tag{11.2}$$

where

$$\left.\begin{aligned} \xi_w'' \cos \delta_s'' &= \xi_w \cos \delta_s - \xi_w' \cos \delta_s' \\ \xi_w'' \sin \delta_s'' &= \xi_w \sin \delta_s - \xi_w' \sin \delta_s' \end{aligned}\right\}. \tag{11.3}$$

The value of D_s'' may be taken as O − P and the analysis for ξ_w'' and δ_s'' is carried out exactly as described in § 9, except that

$$\sigma_0 D_{s0}'' = 29\varepsilon_0'' + 29c'' (-1)^s / R_0. \tag{11.4}$$

It is possible to use the first approximations for predictions for later years and to use values of (O — P) from later observations, in which case we count L from the same zero value used in the original analysis.

12. *Large Diurnal Constituents and Constant Currents.*

If there are very large diurnal currents then our analysis may fail, because M_2 is no longer "predominant." Successful analyses have been carried out for the St. of Canso, Canada, where sometimes the current does not reverse owing to large diurnal currents as well as a constant current.

In such a case the sequences of times of slack water were incomplete, and arbitrary large values were placed in the gaps, for the purpose of the first approximation. Then with the second approximation the values of $O-P$ can be assumed to be zero. A third approximation may be necessary, as the constants are rather susceptible to changes in the value of C.

It should specially be noted that it is well to diminish C for prediction, as if it is too large there will be many gaps in the predictions where no gaps occur in the observations.

13. *Daylight Observations.*

When observations are only possible in daylight each sequence is continuous for about a fortnight and then observations are lacking for about another fortnight. The sequences are not continuous together. In such circumstances it is almost necessary to have observations taken over 12 or 13 consecutive months.

A method of analysis has been devised at the Tidal Institute on the lines of the method of analysis used with hourly heights, as described in the paper referred to in § 1. Any method of analysis will involve corrections for one constituent upon another, and in the following exposition, the method used in the Institute is abbreviated. Much labour has necessarily been required to determine the effects of all the processes upon the different constituents, but it is deemed unnecessary to enter into detailed explanations as the methods are very similar to those of the paper quoted.

Daily multipliers for 15 days are taken as follows :—

d_0, d_1 : unity throughout, $L-\bar{L}=-7$ to 7,

d_a, d_b : -1 for $L-\bar{L}=-7$ to -1, zero for $L=\bar{L}$,
$+1$ for $L-\bar{L}=1$ to 7.

d_2 : $+1$ for $L-\bar{L}=-3$ to 3 inclusive and otherwise -1.

It is necessary to choose the mean values ($\bar{L}$) of L in each fortnight according to a definite rule; we take

$$\bar{L} = -7s \pm (14, 43, 71, 100, 128, 157).$$

Values of D_{s0}, D_{sa}, D_{s2} are required for each fortnight, noting that $D_{sb} = D_{sa}$, $D_{s1} = D_{s0}$.

If we combine the 12 values of D_{s1} from the 12 fortnights with the 12 multipliers m_1 given below, so that we take $\sum_{L=-157}^{157} m_1 D_{s1}$, we obtain the quantity D_{s11}; similarly we require to compute D_{s00}, D_{s11}, D_{s1a}, D_{sa1}, D_{saa}, D_{s20}, D_{s22}, D_{s2b}, D_{sb0}, D_{sb2}, D_{sbb}.

$$\bar{L} + 7s.$$

Multiplier	−157	−128	−100	−71	−43	−14	14	43	71	100	128	157
m_0	1	1	1	1	1	1	1	1	1	1	1	1
m_1	−2	−1	−1	1	1	2	2	1	1	−1	−1	−2
m_a	−1	−2	−2	−2	−2	−1	1	2	2	2	2	1
m_2	1	0	−1	−1	0	1	1	0	−1	−1	0	1
m_b	1	2	1	−1	−2	−1	1	2	1	−1	−2	−1

We then proceed to combine these functions as follows, in order to isolate one constituent from another:—

$$A_{s1} = D_{s11} + 0{\cdot}62\, D_{saa} - 0{\cdot}0025\, D_{s00} - 0{\cdot}037\, D_{s20}$$

$$B_{s1} = D_{s11} - 0{\cdot}72\, D_{saa} + 0{\cdot}0029\, D_{s00} + 0{\cdot}043\, D_{s20}$$

$$A_{sa} = D_{sa1} - 0{\cdot}88\, D_{s1a} - 0{\cdot}20\, D_{sb2}$$

$$B_{sa} = D_{sa1} + 0{\cdot}71\, D_{s1a} - 0{\cdot}14\, D_{sb2}$$

$$\alpha_{11} = A_{21} + A_{01} \qquad \alpha_{1a} = A_{2a} + A_{0a}$$

$$\alpha_{11}' = A_{21} - A_{01} \qquad \alpha_{1a}' = A_{2a} - A_{0a}$$

$$\alpha_{21} = A_{31} + A_{11} \qquad \alpha_{2a} = A_{3a} + A_{1a}$$

$$\alpha_{21}' = A_{31} - A_{11} \qquad \alpha_{2a}' = A_{3a} - A_{1a}$$

Similarly we have

$$\beta_{11} = B_{21} + B_{01}, \qquad \beta_{1a} = B_{2a} + B_{0a}, \quad \text{etc.}$$

We then have

$$\left.\begin{aligned} \alpha_{11} - 0\cdot16\,\alpha_{1a}' &= 801\ \xi_1 \cos(\delta_1 - \Delta_1 + 7\mu) \\ \alpha_1 + 0\cdot23\ \alpha_{11}' &= 956\ \xi_1 \sin(\delta_1 - \Delta_1 + 7\mu) \end{aligned}\right\} P_1$$

$$\left.\begin{aligned} -\alpha_{1a}' - 0\cdot25\,\alpha_{11} &= 945\,\xi_1 \cos(\delta_1 - \Delta_1 + 7\mu) \\ -\alpha_{11}' + 0\cdot17\,\alpha_{1a} &= 780\,\xi_1 \sin(\delta_1 - \Delta_1 + 7\mu) \end{aligned}\right\} \nu_2$$

$$\left.\begin{aligned} \beta_{11} + 0\cdot02\ \beta_{1a}' &= 790\ \xi_1 \cos(\delta_1 - \Delta_1 + 7\mu) \\ \beta_{1a} - 0\cdot02\ \beta_{11}' &= 833\ \xi_1 \sin(\delta_1 - \Delta_1 + 7\mu) \end{aligned}\right\} K_1 \text{ (and } O_1\text{)}$$

$$\left.\begin{aligned} -\beta_{1a}' + 0\cdot04\ \beta_{11} &= 802\ \xi_1 \cos(\delta_1 - \Delta_1 + 7\mu) \\ -\beta_{11}' - 0\cdot03\ \beta_{1a} &= 796\ \xi_1 \sin(\delta_1 - \Delta_1 + 7\mu) \end{aligned}\right\} N_2$$

A similar set of formulæ results when α, α' have suffix 21 or $2a$, ξ, δ, Δ have suffix 2 and 7μ is replaced by 14μ.

We also have

$$\begin{aligned} A_{s22} &= D_{s22} + 0\cdot01\,D_{s11}, \\ A_{sbb} &= D_{sbb} + 0\cdot02\,D_{s20} + 0\cdot05\,D_{saa} + 0\cdot0013\,D_{s00}, \\ A_{sb2} &= D_{sb2} + 0\cdot02\,D_{s1a}, \\ A_{s2b} &= D_{s2b} + 0\cdot02\,D_{sa1}, \end{aligned}$$

whence

$$\left.\begin{aligned} -A_{022} - A_{222} + A_{0bb} + A_{2bb} &= 646\,\xi_1 \cos(\delta_1 - \Delta_1 + 7\mu) \\ -A_{0b2} - A_{2b2} - A_{02b} - A_{22b} &= 693\,\xi_1 \sin(\delta_1 - \Delta_1 + 7\mu) \end{aligned}\right\} K_2,$$

$$\left.\begin{aligned} A_{02b} - A_{22b} + A_{0b2} - A_{2b2} &= 671\,\xi_1 \cos(\delta_1 - \Delta_1 + 7\mu) \\ -A_{022} + A_{222} + A_{0bb} - A_{2bb} &= 621\,\xi_1 \sin(\delta_1 - \Delta_1 + 7\mu) \end{aligned}\right\} J_1 \text{ (and } Q_1\text{)}.$$

A similar set of formulæ arise when the first suffixes of A are taken as 1, 3 instead of 0, 2, and ξ, Δ, δ take suffix 2, while 7μ is replaced by 14μ.

Also

$$\left.\begin{aligned} D_{s20} + 0\cdot0667\,D_{s00} &= 234\xi_s \cos(\delta_s - \Delta_s + 7s\mu) \\ D_{sb0} - 0\cdot05\quad D_{sa1} &= 212\xi_s \sin(\delta_s - \Delta_s + 7s\mu) \end{aligned}\right\} S_2,$$

and the computation of ϵ_0 and C follow from

$$D_{s00} - 0\cdot05D_{s11} + 0\cdot07D_{s20} = 621\,(\epsilon_0 + s180^\circ + 90^\circ) + 356\,C\,(-1)^s/R_0.$$

We thus have four ways of determining ξ_1, δ_1 for S_2 and two ways for each of P_1, ν_2, N_2, ϵ_0, C/R_0, and the results should be combined vectorially.

For K_1 we take

$$\xi \cos \delta = \tfrac{1}{2} (\xi_1 \cos \delta_1 + \xi_2 \cos \delta_2),$$
$$\xi \sin \delta = \tfrac{1}{2} (\xi_1 \sin \delta_1 + \xi_2 \sin \delta_2),$$

and for O_1, using the same quantities

$$\xi \cos \delta = \tfrac{1}{2} (\xi_1 \cos \delta_1 - \xi_2 \cos \delta_2)$$
$$\xi \sin \delta = \tfrac{1}{2} (\xi_2 \sin \delta_2 - \xi_1 \sin \delta_1).$$

Similarly we obtain $\xi \cos \delta$, $\xi \sin \delta$ for J_1 and Q_1 respectively.

The rest of the work is exactly the same as described in the earlier section of this paper; of course, it is not necessary now to correct S_2 for K_2, etc. It is probably advisable to compute second approximations after predictions have been compared with observations.

The Theory of Metallic Corrosion in the Light of Quantitative Measurements.—Part II.

By G. D. Bengough, J. M. Stuart and A. R. Lee.

(Communicated by H. C. H. Carpenter, F.R.S.—Received July 18, 1928.)

[Plate 1.]

In a previous paper* it was shown that if dilute solutions of potassium chloride were allowed to act upon metallic zinc in the presence of oxygen, the corrosion-time curves obtained by means of the observed absorption of oxygen gas were exponential for part of their course. The reason why this particular form of curve was obtained was considered to be the gradual falling off in the concentration of chlorine ions in the experimental conditions of limited volume of solution. For N/10000 and N/5000 solutions the agreement between the experimental and calculated curves was satisfactory up to about 25 days except for a short initial period of two or three days. After about 25 days the experimental curves fell notably below the calculated values, and the suggestion that the reason for this discrepancy was mainly the barring out of chlorine ions from the anodic areas by means of accumulations of corrosion products received some support from the fact that chlorine ions were actually found in solution even after long periods of experiment with the N/5000 solution, in which this discrepancy was most pronounced, but since none were found in the more dilute solutions it seemed probable that some other factor was also operative.

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 451 (1927).

The Production of Hydrogen Gas.

It has been assumed during the early part of the work that no hydrogen gas would be evolved by the action of conductivity water and dilute potassium chloride solutions on zinc, and, in fact, no evolution of gas could usually be seen even in experiments with solutions as strong as N/10. In two or three exceptional cases, notably experiment A 60 with N/5000 KCl, bubbles of gas were seen clinging to the corrosion products, and it was thought that they must consist of hydrogen. To test this, experiment A 60 was stopped at the end of 75 days, and the whole of the gas in the apparatus was forced into an external measuring vessel by means of distilled water; the hydrogen was then determined by burning in the presence of a heated platinum wire. 4·66 c.c. were found. Similar determinations were then made with the gases from several other experiments in which no bubbles of gas had been seen, and the results are given in Table I.

Table I.—Standard-sized Zinc Specimens corroded in 100 c.c. of Solution at 25° C. and 760 mm. pressure.

Specimen.	Concentration of KCl.	Time of experiment. Days.	Hydrogen evolved, cubic centimetres at 25° C. and 760 mm.	Oxygen absorption, cubic centimetres (corrected for hydrogen).	Total corrosion, mgrms. of Zn.	Per cent due to H gas evolution.
A 54	N/20000	62	Nil	7·51	40·18	0
S 1	N/10000	55	Nil	12·68	67·84	0
A 61	N/10000 ($Ba(OH)_2$)	100	Nil	0·18	0·96	0
A 64 (burnished)	N/10,000	64	1·08	5·91	34·49	8·3
A 60	N/5000*	91	4·66	34·65	197·81	6·3
A 63	N/5000	70	2·43	20·63	116·85	5·6
A 59	N/1000	94	8·44	41·01	241·93	9·3
A 62	N/200	75	16·29	75·26	446·13	9·7

Note.—All specimens except A 64 and S1 were annealed for a week at 250° C. Depth of immersion of top horizontal surface 1·5 cm. (not 1·6 cm. as stated by error in previous paper).

* 500-c. c. vessel.

The eye cannot usually detect the evolution of hydrogen gas in the conditions of these experiments even when several cubic centimetres have actually been evolved, and that the total corrosion cannot be deduced from oxygen absorption measurements alone in solutions as strong as, or stronger than, N/10000 KCl, since the percentage of the total corrosion due to hydrogen gas evolution is variable. An exception is experiment S 1, in which "spectroscopically pure" zinc (described in the previous paper) was used.*

Very highly purified zinc is resistant to the action of hydrochloric and sulphuric acids, but the deduction sometimes made that this is due to the prevention of electrolytic action by the elimination of cathodic areas with the impurities in the metal is incorrect because the highly purified zinc corroded as rapidly as the less highly purified Australian electrolytic zinc used in the other experiments, as measured by the absorption of oxygen.† The correct conclusion seems to be that the purer zinc possesses a higher over-potential, which hinders corrosion of the hydrogen-evolution type only.

There seem to be several possible suggestions for the cause of the evolution of hydrogen gas from neutral salt solutions. It may, for instance, be due to direct displacement from water, or to displacement from corrosion products formed by hydrolysis of zinc chloride or from potassium hydroxide formed from potassium chloride.

In support of the hydrolysis hypothesis may be quoted the loss of weight measurements of Bengough and Stuart‡ on the action of zinc chloride on metallic zinc—approximately 107 mgrms. of zinc were corroded by 200 c.c. of a 2 per cent. solution of zinc chloride in 83 days, but no measurements of hydrogen gas were made. To test this, portions of a turned disc of electrolytic zinc were introduced into a eudiometer containing N/50 solution; at the end of 28 days no gas could be observed. This result might possibly be due to the absence of accumulations of impurities of low over-potential, and the experiment was repeated with specimens that had already been corroded in distilled water for 17 days; no gas was observed at the end of 28 days.

In support of the caustic potash hypothesis the fact that fairly strong solutions of alkalis attack zinc with evolution of hydrogen may be quoted. The concentration of alkali will be very low at the beginning of a corrosion experiment, but it will gradually increase as the chlorine ions are replaced by (OH) ions. The concentration will be greatest at cathodic areas and may locally exceed that of the original potassium chloride. Tests made by placing indicators on the surfaces of metal specimens covered with corrosion products at the end of experiments showed strong but irregularly distributed alkalinity; the clear liquid from which the specimen was taken showed approximate neutrality. A test similar to that for the zinc chloride was made to see whether N/100 solution would attack turned electrolytic zinc with production of hydrogen gas, but none was obtained; on repeating the experiment with similar zinc that had been corroded in distilled water, a few small bubbles of gas could be seen at the end of one day, but no further accumulation of gas took place.

---

* The temperature variations in this and the previous paper did not exceed 0·03° C. except when otherwise stated.

† Compare Prost, 'Bull. Soc. Chim. Belge,' vol. 23, p. 94 (1914).

‡ 'J. Inst. Metals,' vol. 28, p. 64 (1923).

It might be thought possible to distinguish between the suggestions by observing the locality at which the gas is evolved. Gas evolution was seen only occasionally and then appeared in the neighbourhood of the pits. Atomic hydrogen is not always produced where the bubbles appear; for instance, if zinc be placed in very dilute acetic acid in a glass dish, bubbles of hydrogen often appear on the floor and walls of the dish, or on minute particles of dust suspended in the liquid*; evidently gas can diffuse away from its cathodic seat of production and appear wherever a suitable nucleus presents itself; possibly the most suitable nuclei are sharp metallic points produced by corrosion.

The question of the production of hydrogen gas is further discussed on p. 93, but it may be stated here the authors are of opinion that the hydrogen is probably formed by reaction with potassium hydroxide.

The experiments summarised in Table I clearly showed the necessity of measuring the amount of hydrogen gas evolved in all future experiments, since its undetected presence causes two errors in attempts to determine the amount of corrosion from oxygen absorption only. These are due to the facts that:—

1. The observed absorption of oxygen is too low, since it is partly replaced by hydrogen.
2. The total corrosion as calculated from the true oxygen absorption is too low by an amount of zinc corresponding to the hydrogen evolved.

It follows that certain curves published in the previous paper, namely, those given for solutions of N/10000 and N/5000 KCl (except for the "spectroscopic grade 'A'" zinc) are merely "apparent oxygen" curves and not true "total corrosion" curves.

---

* Gas bubbles have been seen under the microscope to form on dust particles lying on the metal; these may then be floated up into the liquid by the attached bubbles, which may then be released and the particles dropped on to the metal. Gas bubbles re-form on them, thus producing a periodic evolution of gas.

In Table I of the previous paper a comparison was made between the oxygen-absorption and the loss of weight methods of measuring corrosion. For the latter a correction was introduced to allow for acid attack on the metal during preparation for weighing, and this had the effect of giving a lower result than that ordinarily obtained. In spite of this correction the oxygen-absorption method usually gave a still lower result, and it was stated that "this difference is no doubt partly due to loss of exfoliated metal in the loss of weight method—that losses due to this cause actually took place could usually be seen with the microscope." Since it is now clear that the oxygen-absorption method was itself giving results which were too low, a new set of measurements has been made and is recorded in Table II; in this the true *total corrosion* due to the hydrogen both oxidised and displaced as gas is given.

Table II.—Comparison between Oxygen-Absorption and Loss of Weight Methods of measuring Corrosion.

Specimen.	Concentration of KCl	Zinc by oxygen absorption (corrected for hydrogen gas). mgrms.	Zinc by loss of weight.		Time in acid. minutes.
			Uncorrected mgrms.	Corrected mgrms.	
A 64	N/10000	39·86	39·0	36·4	32
A 65	N/10000	68·64	71·2	63·0	86
A 60	N/5000	197 84	233 0	219·6	118
A 63	N/5000	116 87	129 8	104 2	122
A 66	N/5000	107 27	111 9	104·4	62
A 70	N/5000	76 00	79·0	74 3	44
A 59	N/1000	242 00	262·2	252·2	71
A 62	N/200	446 22	475 6	462·0	41

The oxygen-absorption method still gives results lower than the ordinary uncorrected loss of weight method, but they are sometimes higher and sometimes lower than those given by the corrected loss of weight method. It is clear that over-correction has sometimes occurred, as with specimen A 63. Errors due to exfoliated metal are probably comparatively small since they would tend to raise further the loss of weight figures which are now occasionally higher than the corrected oxygen-absorption figures.

It is evident from the figures given in Table II that though the correction used for the loss of weight method is an improvement, it is not entirely satisfactory, possibly because the nature of the metal surface during the second acid treatment differs from that during the first.

In order to avoid a second acid treatment the gas evolved during the first treatment was collected and measured. The amount of metal corresponding to the volume of gas found was then deducted from the apparent loss of weight. Some results are given in Table IIA.

Some of the above results for oxygen absorption are notably higher than those for the uncorrected loss of weight method. This is a feature that has not appeared before. It is being investigated.

Table IIA.

Specimen No.	Concentration of KCl.	Zinc by oxygen absorption	Zinc by loss of weight. Uncorrected.	Hydrogen evolved in acid treatment (25° C. and 760 mm).	Zinc by loss of weight. Corrected
A 72	N/10000	61·71	60·8	1·3	57·3
A 73	N/5000	85·39	93 0	1·97	87·7
A 68	N/5000 (500 c.c)	294 62	291 0	2·5	284·3
A 74	N/5000 (2 litres)	423 35	386 4	3 94	375·8
A 71	N/1000 (500 c.c.)	574 43	569 4	4 91	556 3
A 69	N/10	2075 64	2140·9	49 15	2009·4
A 76	N/10 (turned)	421 79	453·9	8 c.c.	432·5
A 75	Normal	336 52	369·4	2 51	362·7

Rate of Evolution of Hydrogen Gas.

The method of hydrogen gas measurement described above was unsatisfactory since a determination could only be made at the end of an experiment, whereas it was desired to get curves showing the gas evolution at intervals throughout the experiment. To do this, a thin platinum wire was sealed through the cap of the corrosion vessel (described in the previous paper) and its ends were brought out into glass tubes containing mercury, as shown in fig. 1. A heating current could then be passed through the wire. With N/10000 KCl solution it is possible to determine the hydrogen every third day, and with stronger solutions every other day, but these intervals must be increased in the later stages of experiments.

The method of obtaining the figures used in plotting curves is as follows: "Apparent oxygen" is first measured in the manner stated on p. 437 of the previous paper. The hydrogen gas is then burnt, the remaining gas allowed to cool, and the contraction measured. These operations occupy a period of time, usually from three-quarters of an hour to an hour, during which oxygen absorption is proceeding. Consequently, the oxygen-absorption reading observed next day will be too low by an amount corresponding to the period covered by the hydrogen measurement and a correction must be applied. This correction is taken to be a proportionate amount of that day's apparent oxygen absorption. A typical series of readings is given in Table III, which shows that a corresponding correction must be applied to the apparent hydrogen, which will be too high.

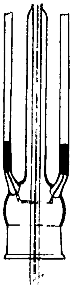


FIG. 1

The rate of evolution of hydrogen gas to be expected from zinc in salt solutions may be considered in conjunction with that from acid solutions, especially hydrochloric and sulphuric acids. This has been studied by many authors—in recent years by Centnerszwer* and his associates, Vondraček and Izakkrizko,† and W. S. Patterson,‡ and the matter has been discussed in detail by U. R. Evans.§ It has been found that the rate of gas evolution for "pure" zinc is constant, but when impurities are present it is often slow at first but increasing with time, especially when the zinc contains impurities with low over-potential such as

Table III.

Days	Oxygen absorbed.		Total hydrogen.	Corrected oxygen absorption.	Oxygen equivalent to hydrogen	Total corrosion referred to oxygen.
	Daily apparent.	Total apparent.				
59·75	1·93*	125·44	38·17*	163·61	19·09	182·70
60·75	1·91*	127·35	39·32*	166·67	19·66	186·33

* These figures have been corrected for the time occupied in burning the hydrogen, as described above.

---

* 'Z. Phys. Chem.,' vol. 87, p. 692 (1914); vol. 92, p. 563 (1918); and vol. 118, p. 415 (1925).

† 'Rec. Trav. Chem.,' vol. 44, p. 376 (1925).

‡ 'J. Soc. Chem. Ind.,' vol. 45, p. 325 (1926).

§ 'Corrosion of Metals,' pp. 64 to 72, 1926 edition (Arnold and Co.).

iron, copper and antimony. Lead at first decreases the rate, but later appears to increase it. The period during which the rate of gas evolution is slow is called the period of induction. The increased rate of evolution of gas is attributed to the accumulations of films of metal of low over-potential, by redeposition or otherwise. Highly purified zinc itself is characterised by a high over-potential; hence most of the common metallic impurities will tend to assist the evolution of gas.

Hedges and Myers* have shown that gas evolution from several metals in acids and alkalis may be periodic in certain conditions, but no corresponding phenomenon has been found in the present research; this may have been due to the relatively long intervals between hydrogen determinations necessitated by the small volumes of gas evolved, or to the absence of the necessary activating agent.

Four hydrogen-time curves are given in fig. 2. That for N/10 KCl, A 69, appears at first sight to be a parabola in the initial stages, but closer investigation shows that it diverges noticeably from this form. On the

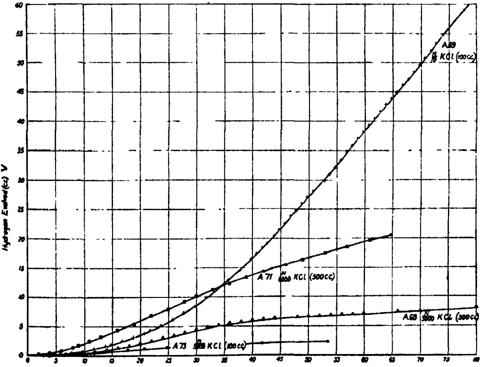


FIG. 2.

---

* 'J. Chem. Soc.,' vol. 125, p. 604 (1924), and vol. 127, p. 445 (1925).

hypothesis that the rate of evolution of hydrogen gas is controlled by the presence of films of metallic impurities of low over-potential, a parabolic form would be expected, but only in certain conditions. For the quantity of impurity produced would be expected to be proportional to the total corrosion, and, therefore, to time *when the corrosion-time curve is a straight line* Also, at any particular instant the rate of evolution of hydrogen would be proportional to the total corrosion and to time when the same condition is fulfilled. Hence, if we represent the volume of hydrogen generated at a particular time as V, the rate of evolution will be dV/dt, and $dV/dt = kt$; hence by integration $V = kt^2$, the equation for a parabola. But for an initial period of about 40 days the required condition is not accurately fulfilled, for reasons discussed later, and a true parabola is not obtained.

The hypothesis can be tested in another way by plotting the rate of hydrogen evolution against corrosion, a procedure which should give a straight line. Such a curve can be obtained from the two curves showing the relation between hydrogen evolution and time and total corrosion and time. It is given in fig. 3 and shows that a straight line is obtained with N/10 KCl, A 69, for a period of about 45 days; for much longer periods divergence would be expected, since hydrogen-gas production must be presumed to be a surface effect, and surface will no longer be proportional to total amount when considerable masses containing impurities have accumulated. In the case of a metal containing large quantities of impurity, the initial parabolic form would probably be passed over very quickly.

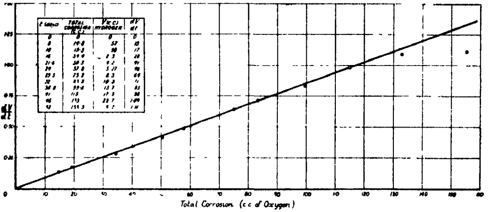


FIG. 3.

Fig. 2 shows that the curves for N/1000 KCl, A 72 (in a vessel with diameter 12·5 c.c.), and N/5000 KCl, A 68, assume a similar form to that for N/10 for the first part of their course, but after a few days the slope rapidly decreases instead of increasing. This is no doubt due to the fact that control of the rate of gas evolution soon passes from the over-potential to some other factor, possibly the lowering of concentration of chlorine ions.

For N/10000 KCl, A 71, the rate of gas evolution was too slow and irregular to plot, and evolution had practically ceased at the end of about 30 days.

It is interesting to note the limits within which the proportion of the total corrosion due to hydrogen gas can vary and the way in which the proportion changes with time and concentration. With very dilute solutions such as N/5000, the corrosion due to hydrogen evolution may be as low as 5 or 6 per cent., as shown in Table I; but with stronger solutions it is usually between 9 and 10 per cent., and may reach 14 per cent. with N/10 at the end of a period of 145 days. During the course of a single experiment the rate of evolution of gas is sometimes maintained for a longer period than the rate of oxygen absorption, so that negative readings are obtained for the latter, as happened in experiment A 60, shown by the curve in fig. 6. It is possible that the relative maintenance of the hydrogen-evolution type of corrosion may be due to the accumulation of KOH towards the end of the experiments, a matter discussed later on, p. 108.

The Relations between Rate of Corrosion and the Oxygen Supply.

It was stated at the end of the previous paper that the factors which were believed to determine the shape of the corrosion-time curves therein published did not include the rate of oxygen supply. It was further stated, " this may be taken to mean that there was always sufficient oxygen available for the fastest rate of corrosion which could occur in the particular conditions studied. The rate of corrosion can be increased by the use of solutions of greater concentration and conductivity, and it is expected that conditions would soon be reached in which the rate of oxygen supply would become the controlling factor. A new branch should then appear in the time-corrosion curves and take the form of a straight line, the slope of which would depend largely upon the rate of oxygen supply." With small supplies of oxygen the slope of this line would be expected to be such that it would be wholly below the exponential curve corresponding to the particular concentration of the solution used; with increased supplies a limiting case would be reached when the slope was identical with that of the maximum of the exponential curve.

These predictions have now been tested for a series of solutions up to normal concentration. Typical curves are shown in figs. 4 and 8. In these experiments the standard sized vessels of 4 4 cm in cross section containing 100 c c of solution were used consequently the maximum possible rate of oxygen supply to the surface of the zinc discs was similar to that in the experiments described in the previous paper

Fig 4 shows that the slope of a curve varies with the concentration of the solution for a given oxygen supply Moreover the curve for A 69 drawn on a smaller scale in fig 9 shows that the slope may increase with time for 40 days or more

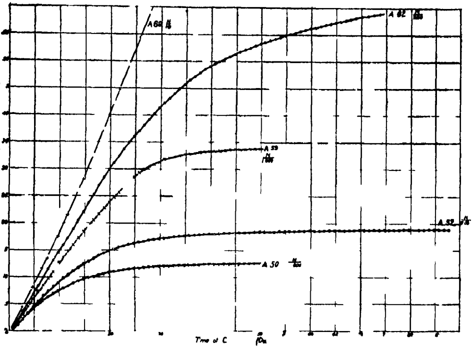
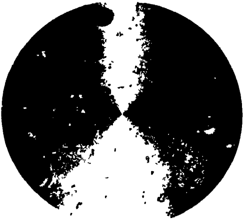


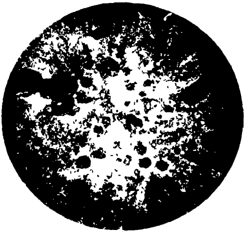
Fig 4

An examination of the surfaces of the zinc specimens showed clearly that with increased concentration an increased proportion of the surfaces had suffered corrosion particularly of the under surfaces which were more widely and deeply attacked than the corresponding upper surfaces (see Plate 1 fig 5) It seems probable that this increase of area attacked with increased concentration is the cause of the steeper slope of the corrosion time curves as the concentration is raised With relatively few pits the corresponding cathodes would only cover part of the available metal surface and the rate of corrosion would be dictated by the rate of oxygen supply to these restricted cathodic areas With more pits the oxygen reaching the correspondingly larger total cathodic

A 52, N 5000 KCl



A 59, N 1000 KCl



A 62, N 200 KCl

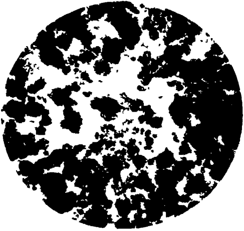


Fig. 5.

area would be utilised, and a greater rate of corrosion would be possible. This rate would reach a maximum when the whole of the metallic surface is being utilised as either anode or cathode, *i.e.*, when the whole of the oxygen that can reach the latter in the conditions of the experiment is being used. This condition is only fulfilled in relatively strong solutions.

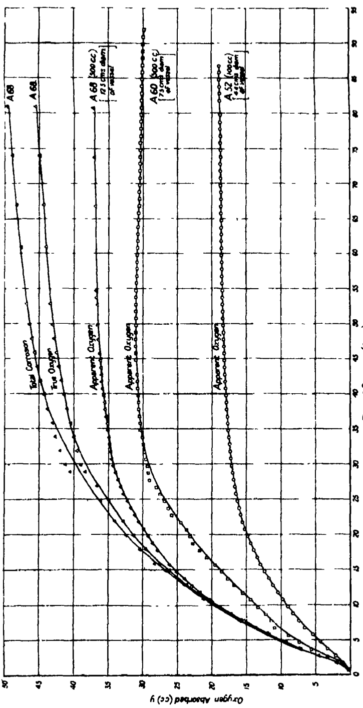
Figs. 4, 8 and 9 show that a maximum rate has been approximately reached with an N/10 KCl solution, but further work is necessary to define its position closely. It seems reasonable to assume that the maximum slope is determined by the maximum possible oxygen supply to the metallic surface in the conditions of the experiment; with more dilute solutions only a part of the maximum possible supply can be used; if this maximum supply be increased, the slope of the corresponding corrosion curve should also be increased.

The oxygen supply is readily increased either by increasing the area of the gas-liquid surface, *i.e.*, by enlarging the cross-section of the vessel, or, for a given cross-section, by decreasing the depth of immersion. Both these methods have been tried. Fig. 6 shows the effect of increasing the cross-section of the corrosion vessel, keeping the concentration constant at N/5000. Fig. 7 shows the analysis of the curve for A 68 in a 12·5-cm. diameter vessel; it is interesting to note that the exponential form is maintained over 70 days, starting from the second day.

This may be compared with the curve for A 52 (also in N/5000 KCl) given on p. 455 of the previous paper, which was described as an exponential. This had an initial slope less than A 68, though a similar one would be expected if both were independent of oxygen control. The explanation seems to be contained in the following figures taken from the experimental results :—

Days of corrosion.	Oxygen absorbed, cubic centimetres.	
	A 52	A 68
2 to 4	2·23	5·1
4 to 6	2·20	4·6

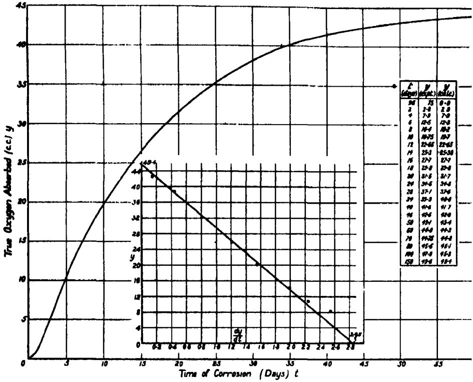
These show that the first part of the curve for A 52 was really a straight line, the slope of which was controlled by the oxygen supply. The exponential curve for the period of 6 to 25 days was determined by the fact that the corrosion which has occurred in the previous 6 days had reduced the concentration to such an extent that corrosion was then controlled by chlorine ions.



A 68
A 68
A 68 (300 cc.) [12.5 cms diam of vessel]
A 60 (300 cc.) [7.5 cms diam of vessel]
A 52 (100 cc.) [4.4 cms diam of vessel]
Total Corrosion
True Oxygen
Apparent Oxygen
Apparent Oxygen
Apparent Oxygen
Time of Corrosion (Days)
Oxygen Absorbed (cc) y
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95

FIG. 6.

It should be stated that experiment A 68 was carried on for 150 days, though only 60 days are plotted on fig. 7. At the end of the experiment the true



t (days)	y (expt.)	y (calc.)
·96	·75	0·9
2	2·8	2·9
4	7·8	7·9
6	12·5	12·8
8	16·4	16·2
10	19·75	19·7
12	22·65	22·65
14	25·3	25·38
16	27·7	27·7
18	29·8	29·8
20	31·5	31·7
24	34·5	34·8
28	37·1	37·6
34	39·9	40·0
40	41·6	41·7
46	42·5	42·9
50	43·1	43·4
60	44·0	44·3
70	44·25	44·8
80	45·5	45·1
100	47·0	45·3
150	49·6	45·4

Fig. 7.

oxygen absorption was 49·6 c.c., or about 4 c.c. higher than the maximum value indicated by the exponential equation. The loss of weight determination after allowing for hydrogen indicated a value about 10 mgrms. (1·9 c.c.) too high. These high values appear to be due to the fact that corrosion-currents have been carried anodically by the (OH) ions of the potash and that part of the corresponding hydrogen has been oxidised. Any corrosion occurring in this way will raise the total amount above the theoretical maximum. This explanation does not account for the fact that the oxygen-absorption method gives a value about 2 c.c. higher than the loss of weight measurement.

It is clear from fig. 6 that an increase in the rate of oxygen supply causes an increased rate of corrosion, but the increase is not directly proportional to the increase in cross-section of the vessel; this would hardly be expected when the cross-section is considerable.

These facts raise the question of the best-sized vessel to use in corrosion experiments. Theoretically vessels with a cross-section of only a few centimetres will restrict the rate of oxygen supply to the metal surface, but in distilled water and very dilute solutions this restriction will not be felt, and an exponential curve will be obtained. With stronger solutions it becomes appreciable and a larger vessel must be used to get this type of curve. Beyond a certain size the restriction by the vessel will be negligibly small.* This size can be determined experimentally and it has been found that an increase of diameter from 12·5 to 15 cm.† does not appreciably increase the rate of corrosion of the standard-sized zinc specimens tested in N/5000 KCl solution in the usual conditions of the present series of experiments. That is to say, that in order to avoid oxygen control the minimum diameter of a vessel for zinc specimens of standard size in N/5000 KCl solution is 12·5 cm.; and a vessel at least as large as this should be used for all stronger solutions. If it is not convenient to use a vessel large enough to give this maximum rate of corrosion, then it is necessary to state the cross-section of the vessel used and the depth of immersion of the specimens, otherwise the experiments cannot fully be interpreted or even reproduced; previous authors have usually been content to state only the volume of liquid in which their experiments have been conducted, and often omit to state whether precautions have been taken to shield the surfaces of the corroding liquid from draughts and to limit convection currents arising in other ways.

It is interesting to consider how far the oxygen supply to the specimens in the present research is dependent upon diffusion. If it is due to diffusion only, it ought to be feasible to calculate the maximum possible rate of corrosion due to oxygen absorption in the given experimental conditions.

In the case of experiment A 69, the specimen of zinc (2·54 cm. diameter, 0·6 cm. thick) was placed at a mean depth of 1·8 cm. in 100 c.c. of N/10 KCl solution contained in a vessel of 4·4 cm. diameter.

The area of the specimen was 14·9 sq. cm., which is approximately equal to the area of the cross-section of the vessel. Hence, for a first approximation we may consider the steady diffusion of oxygen through a cylinder of liquid 1·8 cm. long, 14·9 sq. cm. cross-section, with a pressure of 750 mm. at one end and a zero pressure at the other.

According to Hüfner‡ the coefficient of diffusion of oxygen (K) through water at 25° C. is 1·98, *i.e.*, K = mass of gas passing through unit cross-section in 24 hours with unit concentration gradient.

---

* Even in this case an exponential curve may not be obtained with an oxygen pressure of 760 mm. with solutions between about N/1000 and N/10.

† The 15-cm. vessel held 2 litres of solution.

‡ 'Ann. Phys. Chem.,' vol. 60, p 134 (1897).

Therefore volume of gas flowing through the cylinder in 24 hours will be given by $V = KA\, dc/dx$, where dc/dx is the concentration gradient, c being expressed as volume of gas in unit volume of liquid, and A is the cross-section of the cylinder.

$$dc/dx = 0 \cdot 03/1 \cdot 8.$$

(Volume of oxygen in 1 c.c. of liquid at surface = 0·03 c.c.)
Therefore

$$V = 1 \cdot 98 \times 14 \cdot 9 \times 0 \cdot 03/1 \cdot 8 = 0 \cdot 49 \text{ c.c.}$$

Hence a maximum absorption of 0·49 c.c. of oxygen per day would be expected. Actually the observed maximum absorption was 3·08 c.c

This discrepancy may be explained by the suggestion made by Hufner, who showed that it was not possible to measure the true diffusion coefficient by experiments in which the gas was caused to diffuse through the liquid from the top to the bottom, because downward streaming occurs; this Hüfner considered to be due to the gas dissolving in the liquid and increasing the density of the surface layers.

In experiments where the liquid is exposed to the air the effects due to evaporation* will occur in addition to the above and will be much greater, but it seems clear that even in the present set of experiments, in which convection is reduced to a minimum, most of the oxygen is still conveyed to the metal specimens by this process rather than by diffusion. That the convection currents are very mild is evident from observations made on the distribution of precipitated corrosion products. For instance, circles of zinc hydroxide are often seen round corrosion centres; whereas a difference of temperature of 0·1° C. between the two sides of the containing vessel will set up currents vigorous enough to arrange them in the form of open " horse-shoes "

Variation of Corrosion Rate in Strong Solutions.

The curve for N/10 KCl, A 69, shown in fig. 4 and on a smaller scale in fig. 9 presents a feature not shown by curves for more dilute solutions. The slope of the oxygen-absorption curve increases with time from 1·84 to 3·08, an increase of 1·67 times the initial slope, and this maximum slope is only obtained after about 45 days, whereas usually the maximum is reached in about 3 or 4 days. Moreover, the maximum slope is maintained for a period of 50 days, in spite of the great accumulation of corrosion products. The reason for the increase of slope with time is not obvious. It is difficult to visualise any change at the anodes that would increase corrosion with time, for any increase in the e.m.f. or conductivity of the cells, such as an enlargement of anodic areas, would not increase the rate of corrosion because the process is controlled by the rate of oxygen supply. It is possible that cathodic efficiency may increase with time owing to actions similar to those observed by Hinshelwood* and Palmer† for the oxidation and reduction of copper in the presence of hydrogen and oxygen, and by Constable,‡ who has shown that the surface area of electro-

---

* Adeney, 'Sci. Proc. Dub. Soc.,' vol. 16, p. 143 (1920).

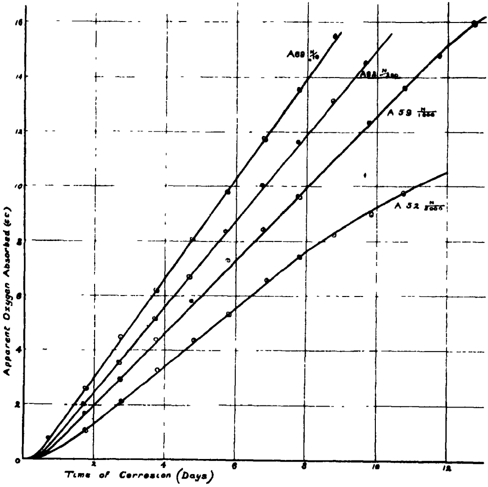


FIG. 8.

---

* 'Roy. Soc. Proc.,' A, vol. 102, p. 318 (1922).

† 'Roy. Soc. Proc.,' A, vol. 103, p. 444 (1923).

‡ 'Nature,' vol. 120, p. 769 (1927).

lytically deposited nickel may be considerably increased by alternate oxidation and reduction; also that the true metallic area may be 1·84 times the apparent area. In the present experiment it seems possible that a gradually thickening layer of *porous* zinc characterised by large surface area may be formed at the

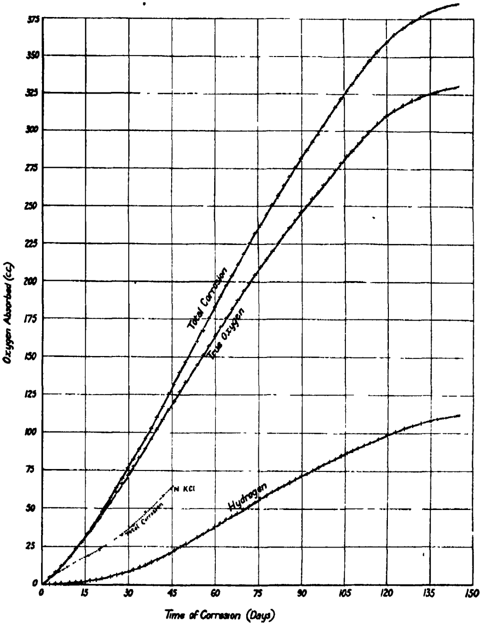


FIG. 9.

cathodes which would be easily oxidised and consequently give a gradually increasing rate of depolarisation; but the layer would not be expected to increase in thickness indefinitely, and when it reached its maximum thickness the rate of depolarisation and consequently of oxygen absorption would remain constant. An increase in corrosion rate during the course of an experiment has also been found with A 75 in normal KCl solution.

An alternative explanation of the increased rate of corrosion is that it is due to the gradually changing surface of the metal as corrosion proceeds. If the total surface of the specimen be increased, it must be presumed that the effective cathodic area will also be increased; for examination of the surface of specimens corroded on strong solutions shows clearly that the original anodic and cathodic areas have undergone considerable interchange. The highly porous nature of the corroded surface may be a factor which increases the efficiency of cathodic depolarisation.

It will be seen from fig. 4 that the slopes of the curves do not increase in direct proportion with the concentration, for instance, an increase of concentration from N/200 to N/10, *i.e.*, 20 times, hardly increases the slope to a greater extent than an increase of concentration from N/1000 to N/200, or N/5000 to N/1000. In fig. 11 rates of corrosion are plotted against the logarithms of the concentrations. These rates of corrosion are obtained

(1) With oxygen-controlled experiments, from the slope of the straight-line portions of the corrosion-time curve.

(2) With experiments controlled by chlorine-ion concentration from the maximum slope of the exponential curve.

This maximum slope of the exponential curve is a measure of the *characteristic rate of corrosion* of annealed zinc in the experimental conditions of concentration and temperature. This rate would be maintained continuously if the concentration of chlorine ions could also be maintained at its initial value.

The rate of corrosion does not increase continuously with concentration, but falls off appreciably in strong solutions, a fact already recorded by Heyn and Bauer, Friend and others, in experiments of the usual type. The amount of precipitated corrosion product falls off, and with a saturated solution of potassium chloride little or none can be found. This is probably due, largely, to the fact that complex salts are formed with zinc in the anion.

The results already described may be briefly summarised as follows for the ranges of conditions covered by the experiments :—

(1) In a solution of given concentration, the rate of corrosion is determined by the effective oxygen supply till this reaches a certain value beyond which it has no further effect. This value is that which permits the time-corrosion curve to take exponential form, it varies with the concentration.

(2) The concentration of the solution and the physical condition of the cathodes largely decide what proportion of the maximum oxygen supply permitted by the experimental conditions actually becomes effective.

(3) The characteristic rate of corrosion of zinc is given by the maximum slope of the exponential curve obtained in specified conditions of concentration of chlorine ions and temperature.

(4) With very strong solutions, such as those near saturation, the rate of corrosion falls with increase of concentration.

Factors which dictate Total Corrosion.

The question of the factors which dictate the total amount of corrosion deserves some discussion. In this connection a comparison of the curves for A 52 and A 68 is interesting. That for A 52, which was tested in N/5000 KCl in a 4·4 cm. vessel, was given in the previous paper; data for A 68 are given in fig. 7. For A 52 the curve published was actually that for apparent oxygen absorption, but there are various reasons for thinking that relatively little hydrogen gas was evolved and that the curve is a fairly close approximation to the true oxygen curve for three or four weeks from the start. It was exponential in form over a period of about 25 days, and the constant A had a value of 22. After 25 days the exponential curve fell notably below the theoretical curve, probably because the accumulated errors due to lack of hydrogen measurements had become appreciable. The curves for A 68 show that the true oxygen-absorption curve is a true exponential for 70 days, the constant having a value of 45·4. It should also be noticed that the total corrosion curve is not an exponential. Both it and the true oxygen curve indicate a final value for oxygen absorption greater than the theoretical value of 45·4. At the end of 150 days, when the experiment was stopped, the true oxygen absorption was 49·6 c.c. The loss of weight determination (after allowing for hydrogen) indicated a value corresponding to 47·3 c.c.

These facts suggest that part of the corrosion is being carried on by some secondary process independent of the chlorine ions, since the behaviour of these ions determines the form of the exponential. They also suggest that though this secondary process principally affects the production of hydrogen, it also affects the true oxygen absorption. The curves suggest that the process becomes relatively more important as the experiment proceeds. These phenomena can be explained by supposing that they are due to the accumulation of some corrosion product that does not contain chlorine. Such a product is caustic potash, and if corrosion currents have been carried anodically by the (OH) ions of the potash and part of the corresponding hydrogen has been oxidised, the high value indicated by the weight determination is explained.

In Part I of this research, which dealt only with corrosion in conductivity water and small quantities of dilute solutions, it was stated that " hydroxyl ions are considered to be mainly film-formers and only corrosion-current carriers to a relatively slight extent." In the more concentrated and larger volumes of solutions studied in the present research, it is reasonable to assume that the potassium hydroxide may locally reach a sufficiently high concentration at the metal surface (after an experiment has been in progress for some time) to react with zinc and displace hydrogen. A large proportion of this hydrogen would probably appear as gas, since the potash would accumulate mainly at the original cathodes and production of gas would be favoured by the presence of original anodes which would present roughened surfaces protected from oxygen.

Some secondary action of the kind suggested would account for the fact that the production of hydrogen is maintained for longer periods than oxygen absorption, with the result that negative readings for the latter are sometimes obtained in the later stages of the experiment—a fact already noted earlier in this paper. Moreover, the interesting observation has been made that hydrogen was being evolved towards the end of the experiment A 72, when it was found by analysis that no chlorine ions were left in solution.

In this connection experiment A 61 is interesting. This specimen was tested in standard conditions in N/10000 barium hydroxide solution which, of course, contains a large excess of hydroxyl ions. The behaviour of the specimen is shown in the following table :—

Days of corrosion	Total oxygen absorbed	Days of corrosion.	Total oxygen absorbed.
0 to 33	0	0 to 81	0·11
34	0·015	82	0·12
36	0·030	83	0·13
37	0·04	84	0·14
48	0·05	85	0·15
73	0·06	87	0·165
77	0·07	94	0·175
78	0·085	100	0·18
79	0·10		

Total corrosion = 0·86 mgm. zinc.

The total corrosion is approximately the same as in conductivity water, but was not measurable at the end of 33 days. At the end of 100 days the experiment was stopped and the specimen examined under the microscope; it was found that corrosion was only noticeable in the neighbourhood of the glass points of support. The table shows that the corrosion was markedly intermittent. In this experiment the primary anodic product must have been zinc oxide or hydroxide, which would possess little solubility in the conditions used. A film would therefore be formed at the anode and the behaviour of this film would, no doubt, control the course of corrosion. But if potassium hydroxide be substituted for barium, the solubility of the film would be increased and corrosion could proceed.

Comparison of the above experiment with N/10000 barium hydroxide with the experiments with conductivity water (given in the Table IV of the previous paper) show that in all cases the maximum amount of corrosion obtained was approximately 1 mgm. of zinc. As $(OH)'$ is the anion common to all the experiments, it is interesting to ascertain if this amount of corrosion corresponds to what would be expected from a consideration of the solubility of zinc hydroxide. This has been found by Miss Ruth Pirret to be 0·47 mgm. in 100 c.c. of water in presence of oxygen at 25° C. This corresponds to 0·32 mgm. zinc or about one-third of the average amount found in five experiments in conductivity water and N/10000 barium hydroxide. On the assumption that the primary anodic product is hydroxide, about two-thirds of the corrosion that occurs in these liquids must be due to the presence of some anion other than $(OH)'$. It probably can obtain access to the corrosion vessel in the form of dust associated with the surfaces of the metal and glass.

In the previous part of this research it was stated ('Roy. Soc. Proc.,' A, vol. 118 (1927)) that the value of the constant A in the exponential equation "depends mainly upon the initial concentration of chlorine ions but also upon the rate of diminution of their concentration." This statement applied to experiments in vessels of similar volume; when this volume is different the constant would be expected to depend partly on the total number of chlorine ions present and not solely upon the concentration. On the basis of the number of ions present, the value expected for the constant A in experiment A 68 would be $22 \times 5 = 110$, whereas it has been found to be only 45. Since, however, the value depends also upon the rate of diminution of the ions, it is reasonable to seek the cause of the discrepancy in a change of this rate. It was suggested in the previous paper that the solution of the zinc chloride would cause hydrolysis and precipitation of an insoluble oxychloride or "adsorption compound." Hydrolysis will probably take place the more readily in the conditions of A 68, owing to the greater dilution of the zinc chloride solution; moreover, the chances of precipitation of the zinc as zinc hydroxide are probably reduced by dilution. It is possible the chlorine ions may also be withdrawn in other ways. For instance, many authors have stated that zinc chloride can dissolve zinc oxide or hydroxide with precipitation of a "basic salt" on dilution; they may also be locked up in a complex anion with zinc.

If adsorption of chlorine ions by zinc hydroxide occurs to any considerable extent, it would be easy to understand why the rate of withdrawal might vary from one set to another. Adsorption is a surface phenomenon, and any variation in the conditions of precipitation of the hydroxide which affected its state of aggregation might be an important factor in determining the value of A.

In the previous paper the suggestion was made (on p. 461) that an accumulation of corrosion products was responsible for preventing the access of chlorine ions to the anodes during the last stages of an experiment with N/5000 KCl. In recent experiments small amounts of chlorine ions have again been found in the liquid after long periods, when corrosion has nearly ceased, but a general consideration of the whole of the data now available has definitely established that when the corrosion-time curve is a true exponential for the whole of its course, then there are no chlorine ions left in solution at the end of the experiment, *i.e.*, when corrosion has practically ceased.

In many recent experiments the effect of precipitated corrosion products has been found to be unexpectedly small. For instance, the curve for A 69 (fig. 4) shows that it was negligible for 100 days or more though the whole specimen had been buried in a thick white mass since the first few days. The curve for A 68 shows a similar phenomenon. as already stated, the exponential form of curve was maintained for 70 days and the final total corrosion was more than the theoretical value, which suggests that chlorine ions were not appreciably barred out from anodic areas. It now seems clear that precipitated corrosion products have little effect on the rate or total amount of corrosion for about 100 days in the conditions studied, though their effect on distribution of corrosion is, no doubt, important. The explanation of the depression of the experimental curves below the theoretical in the first part of the research is probably the lack of hydrogen measurements.

Influence of some Surface Factors.

In fig. 10 a number of curves are plotted giving the results of further experiments on the effect of conditions at the surface of specimens in N/10000

KCl. It has already been shown that specimens whose surfaces have been turned or emery-ground and tested without further treatment give variable results, and generally much lower than those obtained with turned and annealed specimens. The curve for A 50 may be regarded as a typical curve for apparent oxygen absorption with an annealed specimen. A 64 shows the result for a specimen turned and subsequently burnished with an agate burnishing tool; it is strikingly different from that for A 50, and the total corrosion is much lower. An unusually large volume of hydrogen for so dilute a solution was found at the end of the experiment (see Table I), but even when allowance was made for this the corrected oxygen absorption and total corrosion are only about half the apparent oxygen absorption for A 50.

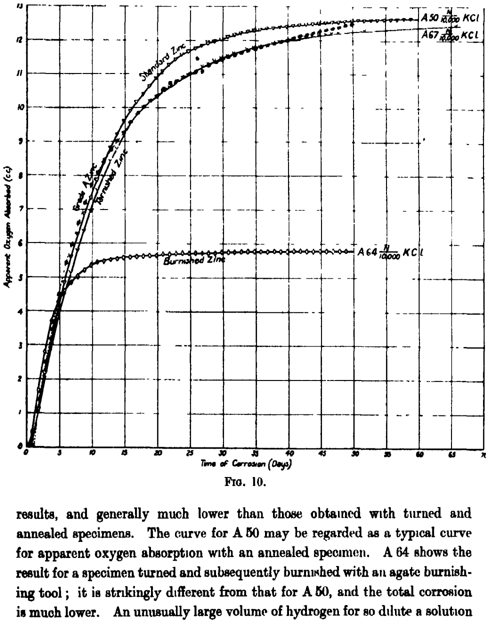


FIG. 10.

It was thought desirable to see whether reduced corrosion of worked surfaces occurred in strong solutions, and fig. 10A shows results for N/10 KCl, specimen A 69 being annealed and A 76 turned. The latter showed a reduced rate of corrosion, thus confirming the earlier results even in conditions in which the corrosion process was controlled by the rate of oxygen supply. The reduced rate of corrosion may be due to the fact that the total number of pits was reduced and consequently the total cathodic area.

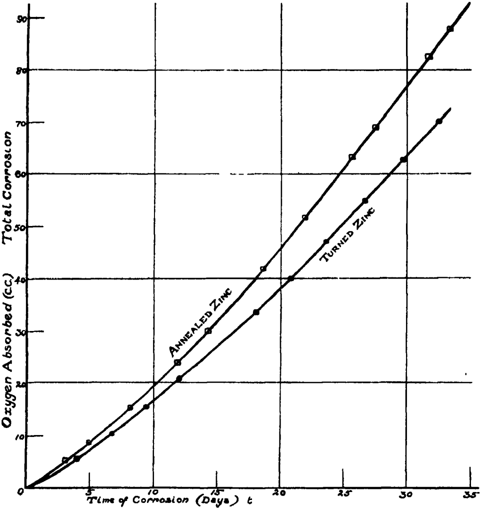
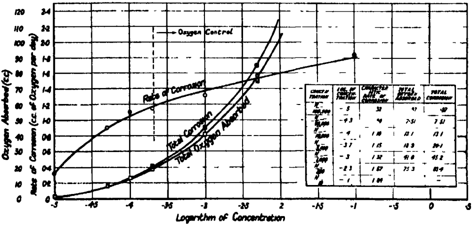


Fig. 10A.

The curve for A 67 is particularly interesting. This specimen acquired a dark purple tarnish while being annealed in argon, doubtless owing to the presence of traces of water vapour or oxygen in the annealing tube. Tarnishes of various colours from yellow to purple had been occasionally observed to form during the annealing process, but all specimens showing them had been rejected owing to possible effects on corrosion. The curve for A 67 shows that the effect has actually been very slight, for it closely resembles that for A 50. This phenomenon seems to be analogous to the neutral effect of the corrosion product in atmospheric corrosion observed by Vernon.



Concentration	Log. of Concentration	Characteristic Rate of Corrosion	Total Oxygen Absorbed	Total Corrosion
N/100,000	-5	32	·47	·87
N/20,000	-4·3	·9	7·51	7·51
N/10,000	-4	1 10	13 1	13 1
N/5,000	-3 7	1 15	18 9	20·1
N/2,000	-3	1 32	41 6	45 2
N/200	-2 3	1 67	75 3	83·4
N/10	-1	1 84		–

Fig. 11.

The curve for grade "A" ("spectroscopically pure") zinc also agrees in form with that for A 50, but the exponential constant A has a value of 11·4 only, that for A 50 being 13·3. A curious feature of the pure zinc curve is that the actual corrosion had exceeded the theoretical maximum by 1·2 c.c. of oxygen by the end of 55 days and would have increased a little farther if the experiment had not been stopped. The last part of the experimental curve showed a fairly constant slope corresponding to an oxygen absorption of about 0·05 c.c. per day, and was quite unlike that of any other curve obtained.

The preparation of this specimen was different from that of the ordinary specimen. The sample from which it was taken arrived from America in the form of rods 0·78 cm. diameter packed in cotton wool in glass tubes closed at the ends with rubber stoppers. The specimen was prepared by breaking off a portion of a rod (by continued bending) of such length (5·7 cm.) that its total surface was approximately the same as that of a standard specimen; it was then placed in the corrosion vessel in a nearly vertical position. The metal was handled only through wash-leather gloves and two thicknesses of filter-paper, and so never came into contact with any metal during preparation for the test. It was not annealed. The exact details of handling in America

Table IV.

Specimen and concentration of solution.	Volume of liquid. c.c.	Oxygen absorbed. c.c.	Initial chlorine in solution. mgms.	Final chlorine in solution. mgms.	Remarks.
A 54, N/20000	100	7·51*	—	Nil	
A 48, N/10000	100	10·95*	—	Trace	
A 50, N/10000	100	13·12*	—	—	
A 67, N/10000	100	12·44*	0·35	Nil	
A 65, N/10000	100	12·83	0·35	0·13	Abnormal curve.
A 64, N/10000	100	6·91†	0·35	0·26	Burnished specimen.
A 72, N/10000	100	10·39†	0·35	Nil	
A 52, N/5000	100	18·87*	0·75	0·13	
A 63, N/5000	100	20·63†	0·71	0·15	
A 60, N/5000	500	34·65†	3·55	0·41	
A 73, N/5000	100	14·69†	0·70	Nil	
A 59, N/1000	100	41·01†	3·50	0·31	
A 71, N/1000	500	97·06†	17·50	3·47	Corrosion still progressing.
A 62, N/200	100	75·26†	17·73	4·55	Corrosion still progressing when experiment was stopped.
A 68, N/5000	500	49·64†	3·50	0·24	
A 74, N/5000	2000	72·47†	14·20	5·10	Corrosion still active when experiment was stopped.
A 69, N/10	100	331·77†	335·00	216·7	,, ,,

* Apparent oxygen absorption.
† True oxygen absorption.

are not known. It is possible that the surface of the specimen had acquired a trace of some soluble impurity which gave rise to an ion that was not gradually removed from the system by precipitation like the chlorine ion; this would enable corrosion to be carried on at a very slow but constant rate. Alternatively an explanation of the excess corrosion similar to that advanced for A 68 may apply in this case also.

Experimental Conditions for estimating Relative Corrodibility.

A question that is raised in an acute form by the shapes of the curves obtained in this research is the period of time to be adopted in carrying out laboratory corrosion experiments designed to assess relative corrodibility. The curves suggest that the best method of estimating relative corrosion would be to continue all experiments till the curves become horizontal, or approximate to this within a defined limit. In strong solutions this necessitates very long experiments.

An alternative method of estimating corrosion, not so satisfactory, but quicker, is to compare the slopes of the curves when oxygen is the controlling factor and consequently the curves are linear. The period of time during which the curves remain linear should also be determined.

The ordinary methods of work based on loss of weight measurements do not permit accurate determination of the forms of the curves, because of the necessity of using large numbers of specimens and of the difficulty of getting strictly comparable and reproducible results. Previous authors have therefore usually selected an arbitrary time of experiment, at the end of which all the specimens have been weighed and compared. The results have been stated in some such form as weight of metal corroded per unit area per day. This procedure would be correct if, up to the end of the period of experiment, all the corrosion-time curves were linear in shape. In the ordinary type of laboratory experiment carried out in small volumes of corroding liquid, the linear form of curve is probably only maintained for long periods with reactive metals in relatively strong solutions.* This gives rise to complications in stating results of experiments in the form of curves in which corrosion is plotted against concentration. Such curves should be derived from corrosion-time curves for various concentrations, and not from data taken at arbitrary times.

Thus if A in fig. 12 represents the true form of curve for a weak solution and B for a strong solution, comparison of the two at time *t* will give quite a different result from that at time *tl*. Moreover, concentration is not the only factor to be considered; volume is also important, for if experiment A be repeated

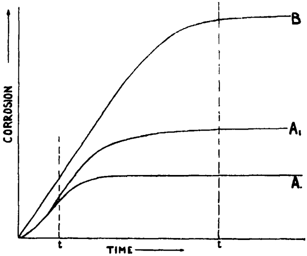


FIG. 12.

---

* The concentration of oxygen gas in such experiments is only partly one-fifth of that in the present set, but this is partly compensated by the much more active convection currents in the former.

with a different volume of solution of similar concentration, a curve A_1 may be obtained lying wholly above A.

The slopes of the curves will be dependent upon the oxygen supply to the metal surface, which is dependent upon

(1) The cross-section of the gas-liquid surface.
(2) The depth of immersion.
(3) Convection.

Of these, (3) is difficult to control or reproduce. In the present research convection currents have been reduced to a minimum (see p. 103).

Another possible plan would be to raise them to a maximum which would be reproducible by some form of vigorous stirring. The drawback to this would be the interference with the distribution of corrosion products, and the extent of the interference would be different with different metals and solutions. Usually some undefined and fluctuating intermediate condition has been adopted which is probably an important cause of lack of reproducibility in corrosion work. It should be noted, however, that reproducibility is artificially favoured by cutting down the oxygen supply to a small fraction of that which could be utilised by the metals under test. Even moderately large differences of reactivity may then be masked; in fact, only metals which corrode so slowly that they cannot make full use of even the restricted supply of oxygen will be clearly distinguished from others.

The above remarks only apply strictly to experiments in which oxygen is the main controlling factor; in certain conditions film formation may, either directly or indirectly, control the corrosion process.

Summary of Results of Parts I and II of this Research.

Typical corrosion-time curves have been defined for annealed zinc in conductivity water and a series of solutions of potassium chloride under one atmosphere of oxygen pressure at 25° C.

These curves usually have a short branch, extending from zero time to two or three days, which is concave upwards. During this period it is probable that a negative electrode potential is being built up at certain isolated positions on the metallic surface till it reaches a maximum fixed by the experimental conditions. This potential is closely connected with the local exclusion of oxygen from the metal surface by the gradual accumulation of precipitated corrosion products. When exclusion is complete, the maximum value is reached, and therefore remains constant.

The short initial branches pass gradually into second branches, which may either be exponential or linear in form, as far as the oxygen-absorption type of corrosion is concerned.

The exponential type is reached in conditions in which the oxygen supply is ample and the main controlling factor is the rate of withdrawal of chlorine ions from the solution. At 760 mm. pressure these conditions only obtain in solutions weaker than about N/1000, even with vessels permitting a large gas-liquid interface and a depth of immersion of 1·5 cm.

The straight-line type is realised in conditions in which oxygen is the controlling factor, *i.e.*, in solutions stronger than about N/1000 in 12·5 cm. vessels. This type is obtained even when hydrogen gas is evolved.

The final branches of the curves are approximately straight lines of low inclination to the horizontal.

With distilled water and solutions of N/20000 or less, little or no hydrogen gas is displaced from the liquid, but with solutions of N/10000 and stronger, hydrogen gas always appears, except when highly purified zinc is used.

The form of the gas-evolution time curves has been determined in certain solutions, and can be explained on the hypothesis that over-potential as affected by metallic impurities is the determining factor.

The proportion of total corrosion due to the gas-evolution type of corrosion varies between about 5 and 15 per cent. for solutions between N/10000 and N/10. The rate of diminution of this type of corrosion is less than that of the oxygen-absorption type.

Curves for total corrosion are not exponential in form, probably because an ion other than the chlorine ion takes part in the corrosion process, especially as regards hydrogen evolution and during the later period of the experiments.

The slopes of the straight-line portions of the curves obtained in oxygen-controlled experiments, *i.e.*, the rates of corrosion, depend not only on the oxygen supply, but also on the concentration of the solution; possible explanations for this phenomenon are discussed.

The total amount of corrosion obtained for standard specimens in water of about 0·055 mho. conductivity is only about 1 mgm. of zinc at the end of 50 days. It is considered that this small amount is due partly to traces of non-metallic impurity, such as dust particles present on the metal surface or on the surface of the vessel.

The total amount of corrosion that has occurred when the corrosion-time curves of KCl solution have become nearly horizontal is independent of the oxygen supply, but is dependent upon

(1) The concentration of the solution.

(2) The rate of withdrawal of chlorine ions from current-carrying capacity either by precipitation, by implication in complex anions, or by some other process.

The effect of precipitated corrosion products on the rate and total amount of corrosion has been found to be very small for a period of as long as 100 days from the start of an experiment.

Corrosion does not increase continuously with concentration but reaches a maximum and then falls off.

The total amount of corrosion in N/10000 KCl is only slightly affected by extreme purification of the metal, this slight effect being due to the suppression of the gas-evolution type of corrosion.

In potassium chloride solutions the total amount of corrosion is greatly affected by the nature of the metallic surface. Highly worked surfaces, *e.g.*, newly turned or burnished, give more variable and usually lower results than annealed surfaces. It has been found that annealing the specimens for a week at temperatures of 250° to 260° C. gives reasonably concordant results; the surface layers of the specimen then show a fine polyhedral structure.

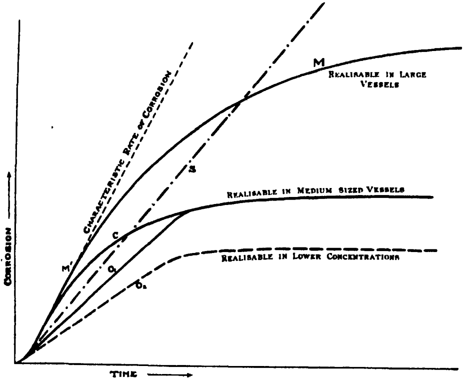


Fig. 13.

Up to this stage of the work the newer version of the electrolytic theory of corrosion, which was outlined in Part I of this research, has stood the test of quantitative investigation; the main result of the research has been to demonstrate which of the several possible factors actually control or notably affect corrosion in specified conditions. The question of the formation of films by the direct oxidation of the metal will be dealt with in a future paper.

A diagrammatic summary of some of the results obtained in the two reports is given in fig. 13, and a general table of data is appended.

This research has been carried out under the direction of the Corrosion of Metals Research Committee of the Department of Scientific and Industrial Research, mainly in the Metallurgical Laboratories of the Royal School of Mines. The authors wish to thank the chairman, Prof. H. C. H. Carpenter, for the many facilities afforded them; also Miss Ruth Pirret, who has given them much valuable assistance.

Types of Curve for Annealed Specimens.

Concentration of solution equals X for curves M, C, and O_1, where X is not greater than N/1000.

M′. Curve characteristic of metal at initial concentration. E/R is the controlling factor. Realisable only with free access of chlorine ions and oxygen. Slope varies with the concentration.

C. and M. Exponential or "available conductivity" curves. "Available conductivity" is the controlling factor in ample oxygen supply. Slope varies with the concentration. M and C result from different volumes of solution. C and O result from different diameters of vessels. O_1 differs from C owing to limited oxygen supply.

S. Total oxygen supply curve for strong solutions, *e.g.*, N/10. Total oxygen supply to metal surface is the controlling factor owing to limited rate of supply. Slope dependent on oxygen supply.

O_1 and O_2. Restricted oxygen supply curves. Controlling factor is the rate of oxygen supply to restricted cathodic areas (restriction due to low concentration of solution). Slope dependent on oxygen supply to restricted cathodes.

With a given value of X, curves M, C, and O_1 are realisable. With a smaller value of X curve O_2 is obtained.

General Table

Specimen No.	Metal sample.	Treatment.	Corroding liquid.	Conductivity of liquid. mhos.	Time of experiment, Days.
A 50	Australian electrolytic zinc	Turned and annealed in argon	KCl N/10000	$15 \cdot 7 \times 10^{-6}$	117
A 52	,, ,,	,, ,,	KCl N/5000	$30 \cdot 6 \times 10^{-6}$	88
A 59	,, ,,	,, ,,	KCl N/1000	$15 \cdot 4 \times 10^{-5}$	94
A 60	,, ,,	,, ,,	KCl N/5000	$31 \cdot 2 \times 10^{-6}$	92
A 61	,, ,,	,, ,,	$Ba(OH_2)$ N/10000	$23 \cdot 15 \times 10^{-6}$	100
A 62	,, ,,	,, ,,	KCl N/200	$73 \cdot 2 \times 10^{-5}$	75
A 63	,, ,,	,, ,,	KCl N/5000	$31 \cdot 0 \times 10^{-6}$	70
A 64	,, ,,	Turned and agate burnished	KCl N/10000	$15 \cdot 5 \times 10^{-6}$	64
A 65	,, ,,	Turned and annealed in argon	KCl N/10000	$15 \cdot 9 \times 10^{-6}$	63
A 66	,, ,,	,, ,,	KCl N/5000	$31 \cdot 4 \times 10^{-6}$	63
A 67	,, ,,	Turned and annealed in argon. Tarnished	KCl N/10000	$15 \cdot 7 \times 10^{-6}$	66
A 68	,, ,,	Turned and annealed in argon	KCl N/5000	$31 \cdot 2 \times 10^{-6}$	151
A 69	,, ,,	,, ,,	KCl N/10	$13 \cdot 8 \times 10^{-3}$	149
A 70	,, ,,	,, ,,	KCl N/5000	$31 \cdot 1 \times 10^{-6}$	25
A 71	,, ,,	,, ,,	KCl N/1000	$15 \cdot 3 \times 10^{-5}$	65
A 72	,, ,,	Turned and annealed in argon	KCl N/10000	$15 \cdot 6 \times 10^{-6}$	47
A 73	,, ,,	Turned and annealed in argon	KCl N/5000	$31 \cdot 3 \times 10^{-6}$	68
A 74	,, ,,	,, ,,	KCl N/5000	—	68
A 75	,, ,,	,, ,,	KCl N	—	45
A 76	,, ,,	Turned	KCl N/10	—	36
S 1	Grade A*	As cast	KCl N/10000	$15 \cdot 2 \times 10^{-6}$	55

* Redistilled zinc supplied by the New Jersey Zinc Company.

of Results.

Total hydrogen evolved c.c.	Total oxygen consumption. c.c.	Total corrosion expressed in oxygen. c.c.	Chlorine in liquid after experiment. mgrs.	Effect of glass supports.	Diameter of corrosion vessel. cm.	Volume of liquid used. c.c.
—	13·12 (app.)	—	—	Noticeable	4·4	100
—	18·87 (app.)	—	0·128	,,	4·4	100
8·44	41·01	45·23	0·315	None	4·4	100
4·66	34·65	36·98	0·41	,,	7·5	500
Nil	0·18	0·18	—	Marked	4·4	100
16·29	75·26	83·41	4·55	None	4·4	100
2·43	20·63	21·85	0·15	,,	4·4	100
1·08	6·91	7·45	0·26	Noticeable	4·4	100
1·71	11·99	12·83	0·13	Not distinguishable	4·4	100
2·29	18·90	20·05	—	,,	4·4	100
—	12·44	—	Nil	Noticeable	4·4	100
10·87	49·64	55·07	0·24	Not distinguishable	12·5	500
112·39	331·77	387·97	216·7	None	4·4	100
1·26	13·59	14·22	—	Marked	4·4	100
20·61	97·06	107·37	3·47	Not distinguishable	12·5	500
1·53	10·39	11·16	Nil	Marked	4·4	100
2·54	14·69	15·96	Nil	Noticeable	4·4	100
13·33	72·47	79·13	5·1	Not distinguishable	15·1	2000
5·07	60·36	62·90	—	Noticeable	4·4	100
16·55	70·57	78·84	—	None	4·4	100
Nil	12·68	12·68	—	—	4·4	100

The Change in Elastic Properties on Replacing the Potassium Atom of Rochelle Salt by the Ammonium Group.

By W. Mandell, Ph.D. (Physics Research Department, King's College, London).

(Communicated by O. W. Richardson, F.R S.—Received July 18, 1928.)

Introduction.

In a former paper the author* made an experimental determination of the elastic properties of crystals of Rochelle salt or Potassium Seignette salt, a double tartrate of sodium and potassium with the formula $Na.K.C_4H_4O_6 4H_2O$. Ammonium Seignette salt ($NaNH_4C_4H_4O_6 4H_2O$) belongs to the same isomorphous group, and can be obtained in large crystals.

It was thought that some interesting results should be obtained on determining the elastic properties of this latter salt, where the ammonium group merely replaces the potassium atom. A comparison of the two results should give some indication of the physical effect due to these chemical changes in the complex molecule.

Method of growing the Crystals.

The crystals of Ammonium Seignette salt were grown in jars in a thermostat. A concentrated solution of the salt was obtained at 28° C.—a higher temperature was not advisable, for some degrees above this temperature the salt in solution began to decompose. Very small crystal "seeds" were fixed by a minute trace of wax on horizontal parallel glass plates which were then placed in the concentrated solution and the temperature so regulated that these small crystals were almost entirely dissolved when the temperature was finally lowered to 28° C. The object of this was to prevent, as far as possible, any unnecessary strains in the crystals during growth. The temperature of the thermostat was then lowered by rather less than one-tenth of a degree per day during the first week, the amounts being gradually increased until room temperature was reached after five or six weeks. The crystals were considerably more difficult to grow than Rochelle salt, as they were especially liable to become coarse and opaque if the temperature was lowered at a quicker rate than this. They are very soluble in water, and plates may be cut easily by means of a wet string.

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 623 (1927).

A Comparison of the Physical Properties of the Two Salts.

Both substances are not very stable. On exposure to the air for some weeks a very thin amorphous white film appears on the surface. This appears in a day or two if the crystals are kept at too high a temperature, and suggests a slight decomposition at the surface, with perhaps some loss of water of crystallisation.

The potassium salt is considerably more dense than the ammonium salt. The density of the former has been found by Buignet[*] to be 1·790, whilst the density of the latter is only 1·587. Besides, the latter is considerably less fragile than Rochelle salt, and is not so liable to fracture due to the warmth of the hand.

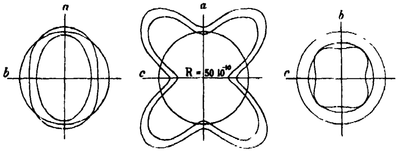


FIG. 1.

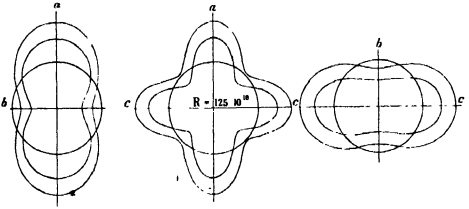


FIG. 2.

The two substances are isomorphous, exhibiting the same class of symmetry and the same facial forms. A. E. H. Tutton[†] has just completed a very large number of physical measurements with isomorphous groups, though not with the Seignette salts. In his general conclusions he found that in any one group there is a small but progressive change in interfacial angular values for the different members, though usually it is only a matter of minutes. Further, he proved that the volumes and edge-dimensions of the space-lattices and the general physical properties of the crystals varied in a progressive manner on substituting certain alkali metals in each of the 18 isomorphous groups with which he dealt. The inter-facial angles for the Seignette salts have been determined by Rammelsberg,[*] and a comparison of the two salts showed differences of less than half a degree between corresponding pairs of faces.

---

[*] 'Journ. d. pharm.,' (3), vol. 40, pp. 161, 337 (1861).

[†] 'Roy. Soc. Proc.,' A, vol. 118, p. 393 (1928).

Ordinary dextro- and lævo-tartaric acids crystallise in identical forms but are enantiomorphous—that is, one is the mirror-image of the other, the dextro variety being distinguished by the right clino-prism {011} only, and the lævo by the left clino-prism $\{0\bar{1}1\}$.

Both of the Seignette salts also appear both as dextro and lævo salts, the former, according to Rammelsberg, being characterised by the presence of the left bisphenoidal face $\{1\bar{1}1\}$, and the latter by the right bisphenoid {111}.

The optical rotation for the optic axial directions was determined for sodium-potassium dextro-tartrate (Rochelle salt) by Pocklington.[†] He found, on using sodium-light, that there was dextro-rotation of 1·2° per millimetre. Dufet[‡] repeated the experiment, and found the slightly larger value 1·35°. For the sodium-ammonium dextro-tartrate, the latter found lævo-rotation for each optic-axial direction of 1·55°. The solutions of both salts, on the other hand, are dextro-rotatory.

The refractive indices have been determined by several experimenters. Lavenir,[§] using sodium-light, found the values for the principal directions for the potassium salt: $\alpha = 1\cdot4900$, $\beta = 1\cdot4920$, $\gamma = 1\cdot4954$, and for the ammonium salt $\alpha = 1\cdot4953$, $\beta = 1\cdot4985$, $\gamma = 1\cdot4996$. Since the refractive indices are so close together, it follows that the double refraction is very feeble, and consequently the dispersion of the optic axes for different wave-lengths is very large. Further, with the potassium salt the optic-axial plane is {010}, and with the ammonium salt, the plane {100}. Sénarmont, Lavenir, Wyrouboff and others[||] have carried out experiments with mixtures of the two salts in "solid solution." They grew crystals containing known proportions of the salts, and examined them in convergent polarised light. It was found that with a gradually increasing proportion of the ammonium salt the optic axes approached each other in a regular manner, and for a certain proportion the uniaxial figure was obtained for a definite wave-length.

---

[*] 'Pogg. Ann. d. Phys.,' vol. 96, p. 18 (1855)

[†] 'Phil. Mag.,' p. 361 (1901).

[‡] 'Journ. d. phys.,' 4 Ser., vol. 3, p. 757 (1904).

[§] 'Bull. Soc. fr. min.,' vol. 17, p. 153 (1894).

[||] 'Zeit. f. Kryst.,' vol. 26, p. 222.

On increasing the proportion the biaxial figure was again obtained, but the axes were now separated in the plane at right angles, the axial angle gradually increasing until the pure salt was obtained.

The experiments showed that the refractive indices varied in an exact linear manner with the composition.

The axial ratios have been determined by Rammelsberg (*loc. cit.*). For the potassium salt he obtained the values

$$a : b : c = 0\cdot8317 : 1 : 0\cdot4296,$$

and for the ammonium salt

$$a : b : c = 0\cdot8233 : 1 : 0\cdot4200,$$

so that the substitution of the ammonium group for the potassium atom causes a slight decrease in the size of the space-lattice in two directions at right angles.

The Elastic Properties of Isomorphous Substances.

It is of interest at this point to examine the elastic properties of isomorphous crystals that have been already determined. Koch[*] found the extension modulus (E) for the cubic crystals rock-salt (NaCl) and sylvine (KCl). He found values of $24\cdot8 . 10^{-10}$ and $24\cdot9 . 10^{-10}$ respectively for directions along the crystal axes. For the directions bisecting two of the axes, and perpendicular to the third, the values were $29\cdot4 . 10^{-10}$ and $47\cdot8 . 10^{-10}$ respectively. Thus, in the directions of the axes, the interchange of atoms caused practically no change in the elasticity, whilst in the other direction the extension was nearly 70 per cent. greater.

The alums are another isomorphous group of the cubic system, but here the molecular structure is very much more complex. For the direction of the axes Beckenkamp[†] obtained the values $55\cdot8 . 10^{-10}$ and $62\cdot3 . 10^{-10}$ for the potassium-aluminium and potassium-chromium alums respectively, and in the other direction, bisecting two axes and perpendicular to a third, values $51\cdot4 . 10^{-10}$ and $56\cdot5 . 10^{-10}$. Thus (E) increases by about 10 per cent. in both directions.

These two examples would suggest that in a salt of simple constitution individual atomic forces may exert a considerable influence on the elasticity in certain directions, while in the more complex substance they have much less influence, and their individual effects are more or less masked by the large number of other atoms in the molecule.

---

[*] 'Wied. Ann.,' vol. 18, p. 325 (1883).

[†] 'Zeit. f. Krist.,' vol. 12, p. 31 (1887).

Experimental Results with Ammonium Seignette Salt.

As stated in the previous paper, for crystals belonging to the Rhombic system, there are nine elastic moduli and nine elastic constants, to determine which, the bending experiments require six suitably cut beams, and the twisting experiments three.

The same apparatus was used as before, and precautions were taken to eliminate knife-edge and other effects which introduced errors.

It was found that the correction due to the " give " of the knife-edge system, together with the " bite " into the beam, was three-tenths of a fringe per load, the number of fringes per load for the different beams ranging from 10 to 18, so that the correction was comparatively small.

For the bending experiments the direction angles which the lengths of the beams made with the crystal axes were slightly different in this experiment, and are given in the following table, together with expressions for each beam, obtained from the general equation for the extension in any direction, viz. :—

$$\mathrm{E} = \frac{\mathrm{I}}{\mathrm{Y}} = s_{11}\alpha_1^4 + s_{22}\alpha_2^4 + s_{33}\alpha_3^4 + (s_{44} + 2s_{23})\,\alpha_2^2\alpha_3^2 + (s_{55} + 2s_{31})\,\alpha_3^2\alpha_1^2 + (s_{66} + 2s_{12})\,\alpha_1^2\alpha_2^2, \qquad (\mathrm{I})$$

where $\alpha_1\ \alpha_2\ \alpha_3$ are the direction cosines of the axis of the beam.

The expressions for the six beams are :—

$$\left.\begin{array}{lllll}
\mathrm{A} & (0 & 90 & 90) & \mathrm{E}_1 = s_{11} \\
\mathrm{B} & (90 & 0 & 90) & \mathrm{E}_2 = s_{22} \\
\mathrm{C} & (90 & 90 & 0) & \mathrm{E}_3 = s_{33} \\
\mathrm{D} & (45 & 90 & 45) & \mathrm{E}_4 = \frac{1}{4}\ [s_{11} + s_{33} + (s_{55} + 2s_{31})] \\
\mathrm{E} & (60 & 30 & 90) & \mathrm{E}_5 = \frac{1}{16}\,[s_{11} + 9s_{22} + 3\,(s_{66} + 2s_{12})] \\
\mathrm{F} & (90 & 45 & 45) & \mathrm{E}_6 = \frac{1}{4}\ [s_{22} + s_{33} + (s_{44} + 2s_{23})]
\end{array}\right\}. \qquad (\mathrm{II})$$

The following table gives the dimensions in centimetres of some of the beams used in the bending experiments, the deflection being the corrected number of fringes with sodium light per load of 58·24 grammes.

An average of at least ten readings was taken for each length of beam, the longest length only being given.

Beam.	Length.	Breadth	Thickness.	Deflection.	E.
A	3·601	1·207	0·1784	18·4	$54·6 . 10^{-10} = E_1$
B	3·428	1·180	0·1854	10·0	$37·7 . 10^{-10} = E_2$
C	4·098	1·067	0·1897	17·1	$36·6 . 10^{-10} = E_3$
D	2·488	0·894	0·1726	15·6	$94·2 . 10^{-10} = E_4$
E	2·803	1·076	0·1579	11·1	$43·2 . 10^{-10} = E_5$
F	3·822	1·018	0·1924	14·3	$37·5 . 10^{-10} = E_6$

Hence from Equation (II) the following values are obtained :—

$$s_{11} = 54·6 \;.\; 10^{-10}, \quad (s_{55} + 2s_{31}) = 285·6 \;.\; 10^{-10},$$

$$s_{22} = 37·7 \;.\; 10^{-10}, \quad (s_{66} + 2s_{12}) = \;\; 99·0 \;.\; 10^{-10},$$

$$s_{33} = 36·6 \;.\; 10^{-10}, \quad (s_{44} + 2s_{23}) = \;\; 75·8 \;.\; 10^{-10}.$$

Thus, on substituting these values in Equation (I), the elastic surface (fig. 1) can be obtained. Any line drawn from the intersection of the axes to these curves gives a numerical measure of the extension modulus (E) for that direction in the crystal.

In figs. 1 and 2 two sets of curves are shown, the inner curve in all cases referring to the potassium salt, and the outer to the ammonium salt, together with a circle of reference (see p. 123).

The Twisting Experiments.

The same apparatus was used as in the experiment with Rochelle salt. A couple of 229·1 gm. cm. was applied to a rectangular beam, and the amount of twist determined for a length of beam, accurately measurable, between two knife-edges which gripped two tinfoil sheets fixed on the crystal. Concave mirrors attached to the knife-edge system gave images of a straight filament on a scale at a distance of 298 cm. from the crystal beam.

In the bending and twisting experiments the results with Ammonium Seignette salt were much more satisfactory than with Rochelle salt.

It was stated in the previous paper that with Rochelle salt, on gradually applying a couple to the beam, a considerable deflection was immediately obtained, followed by a slow continuous movement for a very long period, as if the crystal were gradually breaking down under the continuous shearing forces. This probably explains an effect noticed by Pockels,* who, in determining the piezo-electric moduli of Rochelle salt, found that, on applying a constant pressure in some directions, the electric charge produced gradually increased for some hours. A timing arrangement was again used as in the former experiment, but deflections with the ammonium salt acquired their maximum values almost immediately, excepting for the stretching experiment in the direction of the *a* axis. The constant for the twisting experiment connecting the thickness and breadth of the beam was found to be approximately 0·630, which is the theoretical value for crystals of the Rhombic system.

---

* 'Gott. Abh., p. 183 (1894).

Some of the results obtained were as follows, a reasonable time elapsing between each reading, the deflection being the average of at least ten readings :—

Beam.	Length.	Breadth	Thickness.	Deflection in mm.	T.	Average.
G	3·590	0·945	0·1903	26·0	85·5	$s_{44} = 85 \cdot 7 \cdot 10^{-10}$
G	3·340	0·945	0·1779	29·2	86·2	
G′	2·730	1·021	0·2541	8·4	85·7	
G′	2·670	1·021	0·2127	13·0	85·6	
C	2·830	1·020	0·2467	29·1	355·6	$s_{55} = 352 \cdot 8 \cdot 10^{-10}$
C	2·970	1·020	0·2268	39·1	356·2	
C	2·540	1·020	0·1835	61·8	346·7	
H	1·570	1·143	0·1877	12·2	117·6	$s_{66} = 116 \cdot 1 \cdot 10^{-10}$
H	1·570	1·143	0·1772	14·2	115·8	
H	1·630	1·143	0·1620	18·7	115·0	

On combining these results with those obtained by the bending experiments, values for the elastic moduli are obtained as follows :—

	s_{11}	s_{22}	s_{33}	s_{44}	s_{55}	s_{66}	s_{23}	s_{31}	s_{12}
$10^{-10} \times$	54·6	37·7	36·6	85·7	352·8	116·1	−4·9	−33·6	−8·5

From the relations existing between the elastic moduli and constants a simple calculation shows that the latter have the values :—

	c_{11}	c_{22}	c_{33}	c_{44}	c_{55}	c_{66}	c_{23}	c_{31}	c_{12}
$10^{8} \times$	5·21	3·34	7·63	1·16	0·28	0·86	2·12	5·03	1·83

As with Rochelle salt, the fact that there is no equality between the values c_{44} and c_{23}, c_{55} and c_{31}, c_{66} and c_{12}, respectively, indicates, according to the Cauchy-Poisson theory, that there is a strong polar force between the molecules.

On substituting the above values of the elastic moduli in the general equation for the twisting of a beam of circular cross-section, viz. :—

$$
\begin{aligned}
T_0 = {} & \tfrac{1}{2}\gamma_1{}^4 (s_{55} + s_{66}) + \tfrac{1}{2}\gamma_2{}^4 (s_{66} + s_{44}) + \tfrac{1}{2}\gamma_3{}^4 (s_{44} + s_{55}) \\
& + \gamma_2{}^2\gamma_3{}^2 \{2 (s_{22} + s_{33} - 2s_{23}) + \tfrac{1}{2} (s_{55} + s_{66} - 2s_{44})\} \\
& + \gamma_3{}^2\gamma_1{}^2 \{2 (s_{33} + s_{11} - 2s_{31}) + \tfrac{1}{2} (s_{66} + s_{44} - 2s_{55})\} \\
& + \gamma_1{}^2\gamma_2{}^2 \{2 (s_{11} + s_{22} - 2s_{12}) + \tfrac{1}{2} (s_{44} + s_{55} - 2s_{66})\}.
\end{aligned}
$$

the expression

$$
\begin{aligned}
T_0 = 234{\cdot}4\,\gamma_1{}^4 + 100{\cdot}9\,\gamma_2{}^4 + 219{\cdot}2\,\gamma_3{}^4 + 316{\cdot}9\,\gamma_2{}^2\gamma_3{}^2 \\
+ 64{\cdot}9\,\gamma_3{}^2\gamma_1{}^2 + 321{\cdot}7\,\gamma_1{}^2\gamma_2{}^2
\end{aligned}
$$

is obtained, where $\gamma_1\gamma_2\gamma_3$ are the direction cosines of the axis of the cylinder with respect to the co-ordinate axes.

The outer curves of fig. 2 show the results obtained for the twisted circular cylinder. If the axis of the cylinder coincides with a line drawn in any direction from the intersection of the co-ordinate axes to the surface, then the length of the line gives a numerical measure of T_0 which is the reciprocal of the simple rigidity η.

Discussion of Results.

There is a considerable amount of similarity between the elastic curves of the two salts as shown in figs. 1 and 2. The ammonium salt, both for the twisting and stretching experiments, is less elastic than the potassium salt, so that deformation magnitudes are fairly uniformly increased with the former salt

In conformity with the other changes in physical properties, such as size of space-lattices and of inter-facial angles in the two crystals, the elastic curves show some small differences, especially in the directions where tensions or pressures give a maximum piezo-electric response.

These directions bisect the *ab*, *ac* and *bc* crystal axes of fig. 1. The difference is most noticeable in the *bc* plane, where the curve representing the extension modulus for the ammonium salt almost becomes a circle. It is for pressures in this section that modulus d_{14} for the potassium salt has the very large value, consequently several extra confirmatory measurements were made to check the two elastic curves in this plane.

The general results agree with what is to be expected. The two salts are isomorphous and consequently the two complex molecules must be very similar in structure, and there must be inter-atomic forces of the same nature in the two molecules. Further, it is to be expected that the interchange of the ammonium group for the potassium atom would cause only small differences in elastic properties in any particular direction, for individual atomic forces are largely masked by the large number of other forces in the molecule. Still, it seems somewhat remarkable that the simple chemical change made in such a complex molecule should diminish the elastic properties in all directions of the crystal, although, as stated above, the same thing occurs in the case of the alums which belong to the Cubic System.

It would certainly suggest that the potassium atom must be a sort of " key " atom in the molecule.

In a second paper by the author it will be shown that associated with these chemical and elastic changes, the piezo-electric properties also change, for whereas the value of the modulus for the potassium salt for pressures in the *bc* plane is $1000 . 10^{-8}$, it is reduced to about one-twentieth of the value $(56 . 10^{-8})$ for the ammonium salt

In conclusion the writer wishes to acknowledge his indebtedness to Prof. O. W. Richardson for his very kind interest in this work.

The Determination of the Piezo-Electric Moduli of Ammonium Seignette Salt.

By W. Mandell, Ph.D. (Physics Research Department, King's College, London)

(Communicated by O W Richardson, F R S —Received July 18, 1928)

Introduction.

In a former paper* the author compared the elastic properties of the two Seignette salts. It was shown that these properties were changed considerably on replacing the atom of potassium by the ammonium group. The ammonium salt, when subjected to tensile and torsional forces, was less elastic than the potassium salt, so that its deformation values were considerably increased. Other known differences were referred to, including the sizes of the unit cells, the change in optic-axial angles, with change in composition of the salts in " solid solution," and the dextro- and lævo-rotation of polarised light for the two dextro salts in the directions of the optic axes.

The potassium salt has pronounced piezo-electric properties. When a prism

* *Supra*, p. 122.

is pressed in certain directions with a pressure of one dyne per square centimetre, an electric moment of the abnormally large value of $1000 \, . \, 10^{-8}$ E.S.U. per unit volume is developed.* It was therefore decided to determine the piezo-electric moduli of the ammonium salt and see the effect of the chemical change in the molecule and the associated piezo-electric change.

The General Theory of Piezo-Electricity applied to Crystals of the Rhombic System.

The pyro-electric effect, where charges are developed on heating or cooling certain crystals, had already been observed by a number of observers when P. and J. Curie† discovered that pyro-electric crystals evolved electric charges when they were submitted to mechanically applied pressures. They especially made clear the fact that these charges due to heating are the result of non-uniform elastic deformation, in a manner similar to the electrical effects produced by mechanical tension or pressure. After carrying out experiments which were essentially of a qualitative nature, they deduced a number of important laws which have become the starting point of the theory of piezo-electricity.

Piezo-electricity is essentially associated with effects connected with dielectrics, and equations connecting the electric moment with the associated elastic pressures or deformations were possible from the results of the Curies' experiments, which showed an exact proportionality between the two. This proportionality has been expressed in a linear form by Voigt.

If P_1, P_2, P_3 be the components of the electric moment per unit volume resulting from elastic deformations x_x, y_y, z_z ... x_y, then

$$\left.\begin{aligned} P_1 &= e_{11}x_x + e_{12}y_y + e_{13}z_z + e_{14}y_z + e_{15}z_x + e_{16}x_y \\ P_2 &= e_{21}x_x + e_{22}y_y + e_{23}z_z + e_{24}y_z + e_{25}z_x + e_{26}x_y \\ P_3 &= e_{31}x_x + e_{32}y_y + e_{33}z_z + e_{34}y_z + e_{35}z_x + e_{36}x_y \end{aligned}\right\}, \qquad \text{(I)}$$

where the 18 coefficients e_{ih} are the piezo-electric constants.

On the other hand, if the electric moments be expressed in terms of the six stress components,

$$\left.\begin{aligned} -P_1 &= d_{11}X_x + d_{12}Y_y + d_{13}Z_z + d_{14}Y_z + d_{15}Z_x + d_{16}X_y \\ -P_2 &= d_{21}X_x + d_{22}Y_y + d_{23}Z_z + d_{24}Y_z + d_{25}Z_x + d_{26}X_y \\ -P_3 &= d_{31}X_x + d_{32}Y_y + d_{33}Z_z + d_{34}Y_z + d_{35}Z_x + d_{36}X_y \end{aligned}\right\}, \qquad \text{(II)}$$

where the 18 coefficients d_{jk} are the piezo-electric moduli.

---

* Pockels, 'Gott. Abh.,' p. 183 (1894).

† 'C. R.,' vol. 91, p. 294 (1880).

If for X_x, Y_y, … X_y, and x_x, y_y, … x_y, we write X_1, X_2, … X_6, and x_1, x_2, … x_6, respectively, then, for short,

$$P_j = \sum_h e_{jh} x_h. \quad j = 1, 2, 3. \quad h = 1, 2, \dots 6, \tag{Ia}$$

$$P_j = -\sum_k d_{jk} X_k. \quad j = 1, 2, 3. \quad k = 1, 2, \dots 6. \tag{IIa}$$

The reversible nature of piezo-electric occurrences under isothermal conditions is firmly established by experiment; consequently expressions can be obtained for the thermodynamic potential. In the case of elastic deformations due to mechanical forces under these conditions, it can be shown that the amount of stored energy by work of small amount $\delta\alpha$ is given by

$$\delta\alpha = -\sum_h X_h \, \delta x_h.$$

It is to be noted, however, that under isothermal conditions this is a perfectly general expression; X_h being the forces bringing about the deformation may be either mechanical, electrical or magnetic, so that X_h are generalised forces and x_h functions of X_h the resulting diminution of general co-ordinates. In the case of piezo-electric phenomena we require on the one hand the components of the electric field, and on the other hand the elastic deformation values.

Then the thermodynamic potential takes either of the two forms

$$\xi = -\sum_i \sum_h e_{ih} E_i x_h, \qquad i = 1, 2, 3: \; h = 1, 2, \dots 6, \tag{III}$$

or

$$\xi = \sum_i \sum_h d_{ih} E_i X_h, \tag{IV}$$

where E_i are components of the electric field.

Pyro-electricity is essentially a non-symmetrical phenomenon. Thus a non-directional influence, viz., a temperature change, excites an electric moment with a definite direction. In piezo-electricity also, the idea of stretching is not distinguished by a sign of direction in the one direction more than in the other, so that the phenomenon is associated only with acentric crystals.

In order to obtain the geometrical nature of the piezo-electric constants and moduli as expressed in equations (III) and (IV), certain surfaces are built up from a number of vectors. The variable products $E_i x_h$ and $E_i X_h$ are not components of directed magnitudes, but they can be built up out of several products of such components.

It can then be shown* that equation (IV) is represented by three surfaces, viz.:—

---

* Voigt, 'Lehrbuch der Kristallphysik,' p. 824.

a trivector surface

$$\pm \mathrm{I} = d_{11}x^3 + d_{22}y^3 + d_{33}z^3 + (d_{26} + d_{16})\,x^2y + (d_{31} + d_{15})\,x^2z$$
$$+ (d_{32} + d_{24})\,y^2z + (d_{12} + d_{26})\,y^2x + (d_{13} + d_{35})\,z^2x$$
$$+ (d_{23} + d_{34})\,z^2y + (d_{14} + d_{25} + d_{36})\,xyz,$$

a tensor surface

$$\pm \mathrm{I} = (d_{25} - d_{36})\,x^2 + (d_{36} - d_{14})\,y^2 + (d_{14} - d_{25})\,z^2$$
$$+ 2\,[\{(d_{12} - d_{13}) - \tfrac{1}{2}\,(d_{26} - d_{35})\}\,yz + \{(d_{23} - d_{21}) - \tfrac{1}{2}\,(d_{34} - d_{16})\}\,zx$$
$$+ \{(d_{31} - d_{32}) - \tfrac{1}{2}\,(d_{15} - d_{24})\}\,xy],$$

and a vector, of which the components are

$$2\,(d_{12} + d_{13}) - (d_{26} + d_{35}), \quad 2\,(d_{23} + d_{21}) - (d_{34} + d_{16}),$$
$$\text{and} \quad 2\,(d_{31} + d_{32}) - (d_{15} + d_{24}).$$

These three equations characterise completely the piezo-electric properties of all excitable crystals. In the general case, the hemihedral class of the triclinic system, all the 18 moduli are present, but in other cases the equations simplify very considerably owing to the symmetry. The principle involved is that ξ, as represented by the three surfaces, must be in identical form when referred to a second set of co-ordinate axes allowed by the symmetry.

Or differently expressed, we carry out the covering operations* which belong to the symmetry formula of the particular crystal and still keep the values of the parameters of the surfaces constant.

If, now, the rotation be made about the Z axis the co-ordinates of a point x, y, z become

$$x = x' \cos \theta' + y' \sin \theta', \quad y = -\,x' \sin \theta' + y' \cos \theta', \quad \text{and} \quad z = z'.$$

In the case of Ammonium Seignette salt, there are three two-fold symmetry axes mutually perpendicular to each other; consequently, if the necessary rotation be made about the Z axis, and then about the X axis, it can be shown (Voigt, p. 830) that all the moduli vanish excepting d_{14}, d_{25}, and d_{36}. Thus from equation (II) we get the expressions:

$$\mathrm{P}_1 = -\,d_{14}\mathrm{Y}_z, \quad \mathrm{P}_2 = -\,d_{25}\mathrm{Z}_x, \quad \mathrm{P}_3 = -\,d_{36}\mathrm{X}_y, \qquad (\mathrm{V})$$

so that excitation in Seignette salt takes place only through pressures Y_z, Z_x, X_y, and the moments are then parallel to the X, Y, and Z axes respectively.

---

* Love's 'Elasticity,' 3rd edition, p. 148.

Method of Measurement.

Rectangular slabs, as large as possible, were cut from the crystals and of the required orientations. For the determination of the value of d_{14} for instance, in equation (V) the axis of the length of the slab exactly bisected the angle between the crystal axes b and c, the breadth being in the bc plane, and the thickness parallel to the a axis. Two copper electrodes of the same shape and size were fixed on the two faces in the bc plane, and were slightly smaller than the faces themselves. One electrode was connected to a pair of quadrants of an electrometer, and the other electrode and the second pair of quadrants earthed. Pressures were then applied on the two opposite ends of the slab, *i.e.*, in the directions bisecting the bc symmetry axes, and the deflection θ, due to the developed electric charges, was noted. This deflection was compared with that obtained on earthing one pair of quadrants, and connecting the other pair to two standard cadmium cells (2·0365 volts or $6 \cdot 7883 \, . \, 10^{-3}$ E.S.U.), and noting the deflection θ_0.

If p is the pressure then the moment is equal to $\pm \frac{1}{2} p d_{14}$. If A be the area on which the pressure is applied, a the area of an electrode and C the total capacity of the electrometer, lead wires and crystal slab, then the charge per unit area

$$\pm \tfrac{1}{2} d_{14} \frac{(p)}{\mathrm{A}} = \frac{\theta}{\theta_0} \frac{\mathrm{C}}{a} (6 \cdot 7883 \, . \, 10^{-3}),$$

$$\pm d_{14} = 2 \, \frac{\mathrm{A}}{p} \, \frac{\theta}{\theta_0} \, \frac{\mathrm{C}}{a} (6 \cdot 7883 \, . \, 10^{-3}) = \mu \mathrm{C}.$$

If the units be dynes and centimetres, then the moduli are obtained in absolute units.

The crystal slab stood vertically on a block of hard rubber, a similar block being placed on its upper end. The rubber ensured that the pressure was applied uniformly over the whole of both ends, an experimental detail of importance.

Crystals of the Seignette salts are slightly hygroscopic, so that, under ordinary conditions, the charge would leak away. It was therefore necessary for the experiment to be carried out under the driest conditions obtainable. Consequently the crystal slab was entirely surrounded by a small enclosure containing three removable trays for the drying agent (P_2O_5), whilst the pressure could be exerted from the outside. This enclosure was surrounded by a cylindrical water-jacket to keep a uniform temperature.

The apparatus and the electrometer were arranged as compactly as possible.

The electrometer earthing-key contained a small spiral-spring arrangement, the key being opened and closed by means of a silk thread. This method prevented the collection of static charges on wax or ebonite keys.

Owing to the nature of the crystal, specimens as large as possible were used in order to lessen the loss of charge over the surfaces Even under the favourable conditions under which the experiment was carried out, there was a small leak of charge. For this reason a timing arrangement was used, the initial deflection being obtained by extrapolating for zero time To ensure that the electrometer-needle came to rest as quickly as possible, a thick suspension was used. This also had the advantage that large pressures could be used.

The capacity of the electrometer and lead wires was determined by the method of mixing of charges with a standard condenser The results were then checked by means of a heterodyne wave-meter. The capacity was 29·8 cm.

In carrying out readings one pair of quadrants were insulated and the pressure applied to the crystal The reading was then taken, after which the quadrants were earthed, then insulated, and a similar reading (but of opposite sign) taken on releasing the pressure

A series of readings was taken in this way, time being allowed between observations for the crystal to recover from fatigue. On turning round the crystal slab, pressures were then applied in a direction at right angles on the adjacent faces, for these directions also bisect the crystal axes.

In a similar way suitably cut slabs were employed for finding d_{25} and d_{36} by applying pressures Z_x and X_y and measuring the polarisation along the *b* and *c* axes respectively.

The crystals used in the experiment had been grown about twelve months. Readings were taken for a period extending over several months, and at first they varied in a somewhat capricious manner. A specimen was therefore enclosed for some weeks and it was found that after an exposure to the drying agent for about 24 hours, the values gradually increased for some time and then decreased again to about the normal value. There was also a rather peculiar effect, namely, that the readings differed slightly from day to day about a mean value. The specimens were used under the very driest conditions, and there was the possibility that the drying agent was attacking the water of crystallisation of the salt. It was found, however, that the piezo-electric response had not decreased appreciably after the crystal had been exposed continuously to the (P_2O_5) for three weeks.

In carrying out the actual experiments reasonably consistent results were obtained by preparing the specimens, and then allowing them to recover for a week from the rather violent treatment to which they were subjected during their preparation. Precautions were then taken to ensure that the surfaces, just before use, were free from the white film already referred to. Readings were then taken as soon as the surfaces were dry enough to prevent the charges from escaping over the surfaces.

The following table gives the results obtained for specimens cut from six different crystals. The sign of the charge is inserted in the final values only, the manner in which it is determined being explained below.

Crystal specimen.	Pressure in direction bisecting the axes	Axis along which polarisation takes place.	Load in kilos.	Area of face pressed.	Area of collecting electrode.	$2\theta_0$.	Electrometer deflections (*a*) On adding load. (*b*) On removing load.	$\mu \times 10^8$	Total capacity in cm	Piezo-electric modulus
A 1	+ Y and + Z	*a*	1	1·608 by 0 759	1 350 by 1 052	186	*a* 133 134 135 138 135 135 137 *b* 130 132 133 132 135 133 136 Average 134	1 753	31 1	+53·7 10^{-8}
A 1	− Y and + Z	*a*	1	1 338 by 0 759	1 350 by 1 052	186	*a* 154 154 154 161 161 159 154 *b* 153 153 159 156 156 154 158 Average 156	1 698	31 1	
A 2	+ Y and + Z	*a*	1	1·117 by 0 645	1·350 by 0 875	185	*a* 200 204 202 200 198 203 202 *b* 197 200 198 199 198 197 198 Average 200	1 825	31 1	+58 3 10^{-8}
A 2	− Y and + Z	*a*	1	1 583 by 0 645	1 350 by 0 875	185	*a* 146 146 145 150 150 151 152 *b* 146 145 150 150 148 150 148 Average 149	1 927	31 1	
B 1	+ X and + Z	*b*	1	2 073 by 0·870	1 700 by 0 698	186	*a* 214 218 226 219 224 224 223 *b* 216 222 222 223 222 220 222 Average 221	4 998	30·5	−152 10^{-8}
B 1	− X and + Z	*b*	½	1 079 by 0·870	1 700 by 0 698	186	*a* 204 214 208 212 212 208 210 *b* 214 214 218 213 218 212 212 Average 212	4·988	30·5	

Crystal specimen.	Pressure in direction bisecting the axes.	Axis along which polarisation takes place	Load in kilos.	Area of face pressed.	Area of collecting electrode	$2\theta_0$.	Electrometer deflections (*a*) On adding load. (*b*) On removing load.	$\mu \times 10^6$.	Total capacity in cm.	Piezoelectric modulus.
B 2	+ X and + Z	*b*	1	1·714 by 0·758	1·245 by 0·612	202	*a* 207 211 213 215 214 204 214 *b* 206 204 212 202 204 203 207 Average 208	4·800	30·4	$-147 \, . \, 10^{-8}$
B 2	− X and + Z	*b*	½	1·052 by 0·758	1·245 by 0·612	202	*a* 170 166 167 170 169 166 *b* 170 171 170 166 167 170 Average 168	4·818	30·4	
C 1	+ X and + Y	*c*	2	1·667 by 0·879	1·476 by 1·278	188	*a* 162 168 165 166 163 160 *b* 157 161 162 162 163 161 Average 162	0·928	30·9	$+28·6 \, . \, 10^{-8}$
C 1	− X and + Y	*c*	2	1·664 by 0·879	1·476 by 1·278	188	*a* 159 160 164 164 165 160 *b* 157 161 162 162 163 161 Average 162	0·926	30·9	
C 2	+ X and + Y	*c*	2	1·877 by 0·986	1·550 by 1·550	186	*a* 153 153 160 153 157 155 154 *b* 148 148 149 150 146 151 150 Average 152	0·871	31·1	$+27·9 \, . \, 10^{-8}$
C 2	− X and + Y	*c*	2	1·887 by 0·986	1·550 by 1·550	186	*a* 159 160 156 162 160 164 *b* 157 156 160 156 162 164 Average 160	0·923	31·1	

In applying the small corrections for the capacities of the crystal slabs, the dielectric constants $K_1 = 8·89$, $K_2 = 6·92$, and $K_3 = 6·70$ were used. These are the constants of Rochelle salt; their employment for the ammonium salt will introduce very small errors.

On taking the mean of the above results, we obtain for the values of the moduli $d_{14} = +56·0 \, . \, 10^{-8}$, $d_{25} = -149·5 \, . \, 10^{-8}$, and $d_{36} = +28·3 \, . \, 10^{-8}$. Thus, when centimetre cubes are submitted to unit shears, the units being dynes and centimetres, electric moments are developed whose measures in E.S. units are given by the moduli.

The Determination of the Signs of the Moduli.

The dextro-Seignette salts are characterised by the presence of the left bisphenoidal face {1$\bar{1}$1}, whilst in the lævo variety the right bisphenoid {111} is found instead. The above values refer to the dextro salt, a typical specimen of which is shown in fig. 1. Besides the macro- and brachy-pinacoid, there are three sets of prism faces, whilst the large number of smaller faces bordering on the basal plane are shown for two other typical crystals in the accompanying figures, the bisphenoidal faces being denoted by the letter *o*.

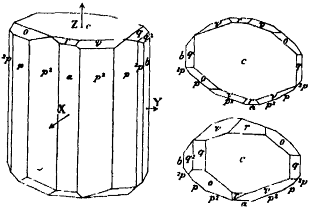


Fig. 1.

Suppose now the + X, + Y, and + Z axes coincide with the *a*, *b* and *c* crystal axes respectively, as shown in the figure, and suppose that a pressure *p* be applied on a suitable specimen in the direction bisecting the + X and + Z axes, and that a positive charge be developed on the electrode on the face in the + Y direction.

Then from equation (V),

$$P_2 = -d_{25}Z_x = -1/2d_{25}p,$$

where in the example given above P is positive, so that the sign of the modulus would be negative.

The Piezo-Electric Constants.

Having obtained the values of the moduli, a comparatively simple calculation is required for the determination of the constants.

On introducing into equation (II*a*) the expression connecting the pressure components and the elastic deformation values, viz. :—

$$X_k = -\sum_h c_{hk}x_h,$$

we obtain the equation,

$$P_j = \sum_k \sum_h d_{jk} c_{hk} x_h,$$

so that on comparing with (I*a*) we get the relation :

$$e_{jk} = \sum_k d_{jk} c_{hk},$$

where the piezo-moduli and elastic constants are now known. On making the necessary calculation the values are as follows .—

$$e_{14} = +6{\cdot}37 \,.\, 10^4, \quad e_{25} = -4{\cdot}11 \,.\, 10^4, \quad \text{and} \quad e_{36} = +2{\cdot}39 \,.\, 10^4$$

Piezo-Electric Response and Temperature.

Valasek,* by submitting specimens of Rochelle salt to pressure and noting the deflections on discharging through a ballistic galvanometer, investigated the variation of piezo-electric response with temperature. Commencing at the temperature of − 75° C., the specimen was allowed to heat up at the rate of one-half degree to one degree per minute. The response was practically nil until a temperature of − 20° C. was reached, when it increased rapidly to a maximum at about 0° C., and then decreased rapidly, the value becoming very small again at about 25° C.

In order to compare the physical properties of the two salts as completely as possible, a similar experiment was carried out with the ammonium salt by the author. Since it would appear that the elastic strains accompanying thermal expansion due to such a rapid rise of temperature must be very great, the rate of increase of temperature was made from one to two degrees per hour, so as to eliminate these strains as far as possible and their associated pyro-electric effects. A quadrant electrometer was used instead of a galvanometer. A simple freezing mixture was employed for the lower temperatures and readings were taken from − 17° C. to + 40° C. It was found that there was no appreciable change between − 17° and + 30°, but on increasing the temperature above this, the crystal gradually became conducting, the charges leaking away in a few minutes. From the magnitude of the first deflection it was evident that the crystal still possessed a considerable piezo-electric response even at the highest temperature for which readings were taken. On allowing the crystal to cool again it gave a somewhat reduced response, which increased in value from day to day.

A small portion of the substance was heated in a test-tube in a water-bath and it was found that just above 30° C. the crystal adhered to the test-tube, indicating a softening at the surface. On increasing the temperature considerably, the substance became viscous, although no ammonia was evolved to indicate decomposition of the salt.

---

* 'Phy. Rev.,' vol. 19, p. 484 (1922).

It would therefore appear that the salt gradually loses its piezo-electric properties owing to this viscous change rather than to any abrupt change in crystalline structure.

Discussion of Results.

The interchange of the ammonium group for the potassium atom causes a considerable change in the piezo-electric properties. The values found above for the ammonium salt are :

$$d_{14} = +56 \cdot 0 \, . \, 10^{-8}, \quad d_{25} = -149 \cdot 5 \, . \, 10^{-8}, \quad \text{and} \quad d_{36} = +28 \cdot 3 \, . \, 10^{-8},$$

whilst the values for the potassium salt are :—

$$d_{14} = +1000 \, . \, 10^{-8}, \quad d_{25} = -165 \, . \, 10^{-8}, \quad \text{and} \quad d_{36}, = +35 \cdot 4 \, . \, 10^{-8}.$$

Thus, for pressures in the direction bisecting the bc crystallographic axes, the polarisation along the a axis decreases by about 95 per cent. on substituting the ammonium group, whilst the other two values are only slightly smaller.

It is also to be noted that the signs of the corresponding moduli are the same. This is of some interest owing to the very close relationship beween piezo-electric phenomena and optical rotation, for the ammonium salt has lævo- and the potassium salt dextro-rotation.

The results show that the moduli are very large. It is therefore somewhat unfortunate that the crystals undergo slight decomposition at the surfaces on prolonged exposure to the atmosphere or to a drying agent.

In conclusion, the writer wishes to acknowledge his indebtedness to Prof. O. W. Richardson for his very kind interest in this work.

The Catalytic Decomposition of Gaseous Acetaldehyde at Surface of Various Metals.

By P. C ALLEN and C. N. HINSHELWOOD.

(Communicated by Sir Harold Hartley, F.R S.—Received July 21, 1928.)

Several investigations have been made with the object of comparing the decomposition of simple substances in a homogeneous gas reaction with the catalytic decomposition which occurs at the surface of a heated metal wire.

Some decompositions, such as that of nitrous oxide and that of hydrogen iodide, which as homogeneous reactions are bimolecular, become unimolecular when catalysed. This behaviour is probably connected with the capacity of metals for absorbing many of the simple gases in the atomic form, and thereby rendering possible a change, such as $N_2O = N_2 + O$, which would require too great an absorption of energy to be possible as a gas reaction.

The homogeneous decomposition of a number of substances with more complicated molecules—such as acetone, acetaldehyde, and other organic compounds—has been investigated, but experiments have not hitherto been made on the kinetics of the corresponding catalytic reactions.

The decomposition of acetaldehyde* is bimolecular as a gas reaction, taking place in accordance with the equation $2CH_3CHO = 2CH_4 + 2CO$. This might have been expected to give place to the unimolecular reaction $CH_3CHO = CH_4 + CO$ at the surface of a catalyst. An examination has therefore been made of the catalytic reaction occurring when the vapour is in contact with heated wires of platinum, platinum-rhodium alloy, gold and tungsten. The results are rather unusual, and the reaction is kinetically of a type not previously encountered.

The apparatus used consisted of a bulb containing an electrically heated wire, the observed resistance of which gave its temperature: the course of the reaction was observed by measuring the increase of pressure on a mercury manometer. The details of manipulation in this kind of experiment have already been described.† Aldehyde vapour was admitted to the reaction vessel at any desired pressure from a bulb of pure liquid aldehyde. The temperature-resistance curves for the actual platinum, platinum-rhodium and gold wires used were specially determined for these experiments. For tungsten Langmuir's data* were used. For all the metals except tungsten, cylindrical reaction vessels were used with the wires stretched along the axis: for tungsten a 60-watt lamp served the purpose very well. A " Hyvac " pump was used for evacuating the apparatus

---

* 'Roy. Soc. Proc.,' A, vol. 111, p. 380 (1926).

† 'J. Chem. Soc.,' vol. 127, p. 327 (1925); 'Roy. Soc. Proc.,' A, vol. 108, p. 213 (1925).

The homogeneous decomposition of acetaldehyde at 500° C. yields methane and carbon monoxide quantitatively, the pressure exactly doubling itself. At higher temperatures Bone and Smith† found that decomposition products of methane also appear. The reaction on the hot wires follows principally the simple course, but with prolonged heating the decomposition products of methane are formed to some extent. With the platinum wire the total pressure increase is barely more than 100 per cent. even after very prolonged heating, and the analysis of the products corresponds almost exactly to pure carbon monoxide and methane With tungsten the increase may reach about 110 per cent, but a definite stage in the reaction appears to be completed in the neighbourhood of 100 per cent., the last 10 per cent apparently being due to a much slower subsequent reaction For the purpose, therefore, of calculating the percentage change from the actual pressure increase at any stage of the reaction, it can be assumed that the "end-point" corresponds to an increase of exactly 100 per cent.

Before the experimental results for the various wires are given in detail it will be well to state one general conclusion. Kinetically the reactions which occur at the surfaces of all four metals are very similar The influence of the initial aldehyde pressure on the rate is the same for each: the heats of activation, though not accurately equal, lie very close together; moreover, the absolute rates of reaction at a given temperature differ very little among themselves. This similarity is most unusual, and might suggest that what is really measured is a homogeneous reaction in a zone of heated gas round the wire. But the measured reaction is not bimolecular: at pressures below about 150 mm. it has a more or less bimolecular character, but at higher pressures it becomes kinetically unimolecular, the change in order taking place exactly as though some " saturation " effect had come into play. Such an effect evidently involves the surface of the wire.

Platinum.

The complete course of the reaction is shown by the following two examples:—

---

* 'Phys. Rev.,' vol. 7, p. 159 (1916).

† 'J. Chem. Soc.,' vol. 87, p. 910 (1905).

Temperature 875° C. Initial pressure of aldehyde = 200 mm.		Temperature 875° C. Initial pressure 200 mm	
Time.	**Increase in pressure.**	**Time**	**Increase in pressure**
Seconds	mm	Seconds.	mm.
0	0	0	0
60	58·5	120	94
120	93	300	143
180	114	480	169
240	131	720	183
360	152	960	192
480	165	1200	194
600	174	1530	198
900	185	1800	200
1500	196	2400	202
2100	200	4080	204

Assumed theoretical "end-point" = 200

The time required for 25 per cent, 50 per cent and 75 per cent. change respectively will be denoted by t_{25}, t_{50} and t_{75}. The ratio t_{75}/t_{50} for these two experiments has the values 2·45 and 2·61 respectively. In a simple reaction of the second order this ratio would be 3, while in one of the first order it would be 2.

The influence of the initial pressure of the aldehyde is such that above 150 mm. t_{25}, t_{50}, etc., are almost constant, while at lower pressures they increase rapidly as the pressure falls.

Temperature 850° C.

Initial pressure (mm.)	25	50	75	100	150	200	250	300
t_{25} (seconds)	480	238	170	118	105	105	95	88
t_{50} (seconds)	1700	650	480	350	275	250	240	230

This fact is clearly illustrated by the first curve in fig. 2, and by the spacing of the curves in fig. 1.

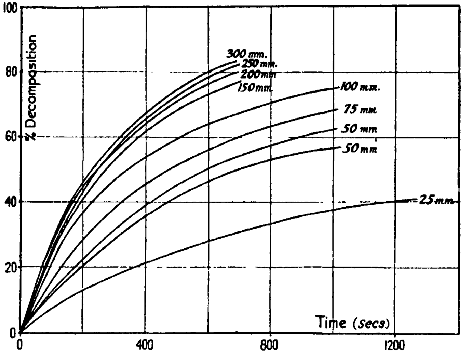


FIG. 1.

The influence of pressure is very nearly the same at all temperatures from 685° to 925° C.

Temperature.	t_{25} (seconds).			t_{50} (seconds).		
	100 mm	200 mm	300 mm	100 mm	200 mm.	300 mm.
° C.						
645	—	—	5940	—	—	—
683	3600	3048	2490	—	—	—
767	660	480	414	2028	1428	1134
793	415	265	242	1560	840	805
815	216	150	144	648	492	432
850	118	105	88	350	250	230
875	62·5	51·5	48·5	177	135·5	125
925	30·5	25	22·5	76·5	58·5	56

The products of reaction exert a certain retarding influence, which, however, is not nearly so marked as, for example, that of oxygen on the decomposition of nitrous oxide at the surface of platinum.

Temperature 875° C.

Initial pressure of aldehyde (mm.).	Initial pressure of products (from previous experiment).	t_{25} (seconds).	t_{50} (seconds).	t_{75} (seconds).
100	0	60	162	396
100	100	75	192	405
100	200	93	219	459

Since it appeared probable that the "poisoning" effect was due to the carbon monoxide contained in the products, a number of experiments were made on the influence of this gas itself.

Temperature 850° C.

Pressure of aldehyde	Pressure of carbon monoxide.	t_{25} (seconds).	t_{50} (seconds).
100	0	115	333
100	50	128	350
100	75	132	370
100	101	150	425
100	208	158	440
100	302	158	395

Several points of interest can be seen from these figures. First, 100 mm. of carbon monoxide reduces the rate in about the same proportion as 200 mm. of the reaction products, which lends support to the supposition mentioned above. Secondly, at a pressure greater than about 100 mm., carbon monoxide ceases to have any further effect, but the rate of reaction is by no means reduced to zero: apparently, therefore, there are various regions of the wire which are able to catalyse the decomposition, only some of which can be poisoned by carbon monoxide.* Thirdly, for small pressures of carbon monoxide, the first additions have rather less effect than subsequent amounts, as becomes evident if the results are plotted on a curve; it is as though adsorption of the carbon monoxide on certain regions of the wire became easier when once a foothold had been gained.

The influence of other foreign gases was examined.

---

* *Cf* Burk, 'Proc. Nat. Acad. Sci.,' vol. 13, p. 67 (1927).

Temperature 850° C. Initial pressure of acetaldehyde 100 mm.

Time (seconds).	Percentage change.		
	No foreign gas.	200 mm. argon.	300 mm. argon.
60	15·8	15	15·5
120	26·1	24·5	26
240	41·4	40	40
420	56·6	55·5	54·5
600	66·8	66	64·5

Temperature 850° C. Initial pressure of aldehyde 100 mm.

Time (seconds).	Percentage change.	
	No foreign gas.	200 mm. helium.
60	14	11
120	25·5	25
240	46	44
420	61·5	62
600	71	73

(The slight difference between the "blank" experiments in this table and the previous one is due to considerable vicissitudes suffered by the wire in many intervening experiments.)

The negative result of these experiments was perhaps to be expected, but they have a certain value as a test of the temperature control, since varying amounts of electrical energy had to be supplied to the wire in different experiments to maintain equal temperatures.

The value of E, the "apparent heat of activation," was calculated from the Arrhenius equation in the usual way, using the values of t_{25} and t_{50} for the experiments with 300 mm. of aldehyde, summarised in one of the previous tables. The 300 mm. values were used because at 300 mm. t_{25} and t_{50} have become independent of pressure. From the slope of the log t_{25}, 1/T curve the result obtained is E = 42,400 calories, while from the corresponding curve for t_{50} the result is E = 47,300 calories. The former value should be very nearly independent of any influence that carbon monoxide exerts on the reaction.

Platinum-Rhodium.

The results obtained with a wire of platinum-rhodium (10 per cent. Rh, 90 per cent. Pt) were generally similar to those found with platinum.

Temperature 883° C.

Initial pressure of aldehyde (mm.).	t_{25} (seconds).	t_{50} (seconds).	t_{75} (seconds).	t_{75}/t_{50}.
25	474	1188	—	-
50	246	670	—	—
100	156	408	1068	2·62
150	128	330	—	—
200	115	306	672	2·20
250	108	294	—	—
300	100	279	645	2·31

The temperature coefficient of the reaction rate was calculated from the following series of results :—

Initial pressure 300 mm.

Temperature.	t_{25} (seconds).	t_{50} (seconds)
° C.		
747	1560	4580
777	804	2280
798	540	1470
832	290	740
863	184	456
883	100	279
926	50·5	123

Values of E, calculated as before from the values of t_{25} and t_{50} respectively, are 45,000 calories and 47,000 calories.

Gold.

With gold, as with the other metals, the "saturation" effect is quite well marked, there being comparatively little change in the value of t_{25} or t_{50} at pressures above 150 mm.

Temperature 758° C. Wire I.

Initial pressure of aldehyde (mm.).	t_{25} (seconds).	t_{50} (seconds)
75	1230	4560
100	720	2370
120	540	1758
150	450	1290
200	420	1182
250	402	1152
300	402	1080

The course of one typical experiment is given below

Temperature 758° C. Wire I. Initial pressure 200 mm.

Time (seconds).	Pressure increase (mm.).	Time (seconds).	Pressure increase (mm.).
60	10	1200	101
180	26	1800	123
360	45	2400	138
600	67	3600	156
900	86		

The ratio t_{75}/t_{50} is again between 2 and 3: at 758° C. with Wire I the ratio was 2·66 for 200 mm. and 2·64 for 300 mm.

The temperature coefficient was determined for two wires. Owing to the short life of fine gold wires which, at high temperatures, are very liable to break under their own weight, the initial rate of reaction was measured rather than t_{25} or t_{50}. Apparatus I did not survive.

Apparatus II.		Apparatus III.	
Temperature.	Initial rate mm./minute for 300 mm. aldehyde.	Temperature.	Initial rate mm./minute for 300 mm. aldehyde.
° C.		° C.	
618	0·80	668	1·15
653	1·27	697	2·70
685	2·21	731	5·1
715	4·20	761	13
743	8·10	789	20
776	19·7		
808	39		
Whence E = 51,000 calories.		Whence E = 49,000 calories.	

The curve log rate, 1/T, for Wire II deviates from a straight line at the two lowest temperatures, which were therefore omitted from consideration in calculating E. The deviation was found to be connected with a certain change in the various adsorptions at the lowest temperatures, which is reflected in a change in the relation between rate and initial pressure. Thus at 618° there is a falling off in the percentage rate with increase of initial pressure, as shown by the following :—

Initial pressure	Percentage change in 70 minutes.
100	30
200	20·5, 22·5
300	15, 16

This is analogous to what happens with nitrous oxide at the surface of platinum, where there is a marked retardation by the products of reaction.

Tungsten.

As has been mentioned, the total pressure increase somewhat exceeds 100 per cent., but the increase beyond 100 appears to be more of the nature of a subsequent decomposition.

Temperature 889° C. Initial pressure 200 mm.

Time (minutes)	1	2	4	6	9	13	20	30	42	60	120	307
Pressure increase (mm.)	23·5	43·5	75·5	99·5	125	152·5	175	195	202	207	214	218

The ratio t_{75}/t_{50} is again between 2 and 3, but nearer to 2.

Initial pressure of aldehyde (mm.) .	100	200	300
t_{75}/t_{50} .. .	2·35	2·03	2·11

"Saturation" with increase of initial pressure is less rapidly reached than with platinum or gold (fig. 2).

Initial pressure (mm.).	t_{25} (seconds).	t_{50} (seconds).
25	264	840
50	228	618
100	159	420
150	156	378
200	145	360
250	132	330
300	108	288

The following results furnish values of E :—

Initial pressure 300 mm.

Temperature.	t_{25} (seconds).	t_{50} (seconds)
° C.		
716	5340	—
756	1740	4410
708	1380	3540
794	708	1980
814	500	1390
838	340	900
857	228	570
876	124	330
889	108	288
	Whence E = 50,000 calories	Whence E = 49,000 calories

Comparison of the Absolute Rates of Reaction.

The measured rate of reaction is directly proportional to the area of the surface of the wire, and inversely proportional to the volume of the reaction vessel. In order, therefore, to make the results for the different wires comparable, the rates for 1000° (abs.) were reduced to what they would have been in vessels of 50 c.c., with wires of 1 square cm. surface area, and with initial aldehyde pressures of 300 mm.

1000° (abs.).	Measured initial rate (mm./minute)	Volume of vessel (c.c.).	Area of surface of wire (sq. cm.).	Reduced rate (mm./minute)	E.
Platinum	5·9	52·5	0·650	9·6	42,400
Platinum-Rhodium	2·13	59·6	0·465	5·5	45,000
Tungsten	1·33	204·0	1·09	5·0	50,000
Gold II	5·5	43·3	0·45	10·5	50,000 av.
Gold III	5·2	40·1	0·44	9·6	

Theoretical Discussion.

The experimental criterion of a " unimolecular " or " bimolecular " reaction is the dependence of t_{25} or t_{50} on the initial pressure ; there are several ways in which a reaction which really depends upon the interaction of two molecules may assume the appearance of a unimolecular reaction, but there is no way in which the converse is possible.

Since, therefore, t_{25} and t_{50} vary markedly with the initial pressure of acetaldehyde, at least over a certain range, we may conclude that the number of molecules involved in a single act of chemical transformation is two and not one.

We have now to consider why the apparent order of the reaction approaches unity for higher initial pressures.

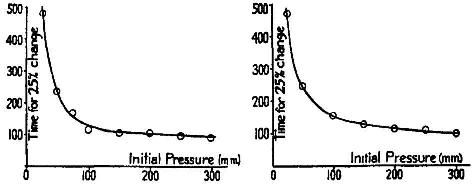
We cannot assume that a unimolecular and a bimolecular reaction occur simultaneously: this would make the predominant order two at higher pressures and one at lower pressures, which is the reverse of the fact. Nor does it help to assume that the bimolecular reaction is "poisoned" at higher pressures, because this assumption would involve an increase in the percentage decomposed in a given time as the initial pressure decreased.

Various ways in which the apparent order of an essentially bimolecular reaction can be reduced must therefore be considered.

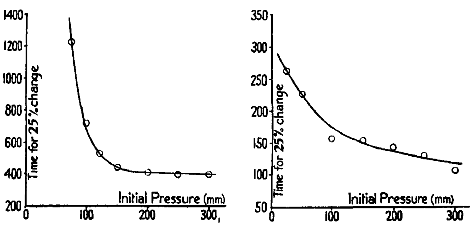
(*a*) If there is a marked adsorption of the products, the apparent order of a reaction may be reduced by as much as a whole unit, because increase of initial pressure involves the presence of an increased amount of the product at every stage of the reaction, and therefore a greater "poisoning" action, which counteracts the increase in rate due to the greater pressure of the reacting substance itself. But if a bimolecular reaction is "poisoned" by its products in this way, markedly enough to appear unimolecular, the ratio t_{75}/t_{50} will be lengthened out to much beyond its normal value 3, whereas in the aldehyde decomposition this ratio is always between 2 and 3. Moreover, direct experiments show that the retarding effect of the reaction products is not nearly marked enough.

(*b*) If the surface of the wire were "saturated" with adsorbed molecules which reacted among themselves, then further increase of pressure could not increase the rate and the reaction would appear to be of "zero" order; if, on the other hand, the number of adsorbed molecules were small, the bimolecular character would reveal itself. There would therefore be an intermediate stage where the reaction would appear of the first order. But this stage could only occur over a limited range of temperature and pressure. The curves in fig. 2 give no indication that at higher pressures the reaction tends toward "zero" order (which would be shown by a minimum in the value of t_{25}). It is just possible that the minimum might be flattened out on account of the existence in the wire of centres of unequal catalytic activity, but from a consideration of all the results this seems improbable.

(*c*) The remaining possibilities involve a reaction between one molecule from the gas and one activated molecule on or from the wire. We might have an activated molecule leaving the wire and reacting by collision with a molecule in the gas: this would be a unimolecular reaction unless the wire were saturated with aldehyde, when it would be of "zero" order: it would only be bimolecular if the pressure were so low that the mean free path were of the same order as the dimensions of the vessel.



Platinum 850° C.   Platinum-Rhodium 883° C.



Gold 758° C.   Tungsten 889° C.

Fig. 2.

If, on the other hand, a molecule from the gas is assumed to react with an activated molecule on the surface of the wire, then the facts can be accounted for: the number of molecules striking the wire is always proportional to the pressure, while the number on the wire is at low pressures proportional to the pressure and at high pressures constant. Thus the reaction is of the second order at low pressures and of the first order at higher pressures. This hypothesis as to the mechanism of the reaction is inherently probable, and will therefore be adopted for the present.

Molecular Statistics of the Reaction

In general the value of E for a heterogeneous reaction is a complex quantity, and depends on the variation of the adsorption with temperature. Only when the adsorption is constant can E be regarded as the energy necessary for reaction. When there is "saturation," as in the mechanism we have assumed for the aldehyde decomposition, this condition may be regarded as satisfied

The number of molecules of acetaldehyde, at 300 mm. pressure, striking 1 square cm. of wire per second is found to be $6 \cdot 95 \times 10^{22}$, when the temperature is taken as that of the cold gas. (Any uncertainty as to the proper temperature to use in the calculation does not, however, affect the order of magnitude.) Let σ be the fraction of the whole surface which, at saturation, is covered with aldehyde molecules. Of these molecules on the wire, a fraction $e^{-E/RT}$ at any instant may be regarded as active, and likely to react if struck by a molecule from the gas, which is not assumed to contribute an appreciable proportion of E, since the gas is cold. If collisions between gas molecules and adsorbed molecules were instantaneous, the greatest possible rate of reaction would be $2 \times 6 \cdot 95 \times 10^{22} \times \sigma \times e^{-E/RT}$ molecules per second. If, however, the collision has a finite duration, the possible rate is greater, because the energy of an adsorbed molecule may change about 10^{11} times per second, this being the frequency of oscillation of the metallic atoms in the wire to which the adsorbed molecules are attached. If the duration of a collision be τ, the possible rate becomes $2 \times 6 \cdot 95 \times 10^{22} \times \sigma \times e^{-E/RT} \times \tau \times 10^{11}$ molecules per second. This is easily reduced to mm./min. for a 50-c.c. vessel.

Using the data in the previous table, the following maximum rates are calculated :—

——	Calculated maximum rate.	Observed rate
Platinum	$2 \cdot 9 \times 10^{8} \times \sigma\tau$	9·6
Platinum-Rhodium	$8 \cdot 0 \times 10^{7} \times \sigma\tau$	5·5
Tungsten	$6 \cdot 6 \times 10^{6} \times \sigma\tau$	5·0
Gold	$6 \cdot 6 \times 10^{6} \times \sigma\tau$	10·0

Since in the homogeneous reaction acetaldehyde reacts at a rate very close to that calculated from the equation—number of molecules reacting = number colliding $\times e^{-E/RT}$—the numbers in the last two columns might be expected to agree also. $\sigma\tau$ would therefore have to be between 10^{-6} and 10^{-8}. If σ approaches unity, the value of τ, the duration of a collision, would be of the order 10^{-6} to 10^{-8} seconds. In this time a freely moving molecule would traverse a large number of molecular diameters: the collision, therefore, cannot be regarded as a perfectly elastic one; in fact, we have a state of affairs similar to the temporary formation of an unstable second layer of aldehyde molecules superposed on the first.

We must now refer again to the surprising fact that reactions at the surface of such different metals as gold, platinum and tungsten should be so similar, not only in regard to dependence upon pressure and temperature, but also in regard to absolute rate. The complete departure from the simple second order precludes the assumption of a purely gaseous reaction near the wire, and the high values of E show that it is not a question of a "diffusion" reaction. But the unusual lack of specificity in the different catalysts suggests that we have to do with adsorption of a feebler, less markedly "chemical" kind, than that involved in catalytic reactions of such gases as nitrous oxide. As regards the activation of the molecules the wire therefore serves mainly as a source of energy, without modifying their stability very profoundly. In accordance with this, the mean value of E for the four catalysts used, 46,800 calories, is almost exactly the same as the value 45,500 calories for the homogeneous reaction.

Summary.

Experiments had previously been made to compare the kinetics of simple homogeneous reactions, such as the decomposition of nitrous oxide, with the corresponding catalysed reactions. In continuation, the catalytic decomposition of a more complex substance, acetaldehyde, the homogeneous decomposition of which is bimolecular, has been investigated.

Electrically heated wires of gold, platinum, platinum-rhodium alloy and tungsten were used as catalysts.

The reaction at the surface of each of these metals involves two molecules of acetaldehyde, but for initial pressures of more than 150 mm. tends to appear unimolecular owing to the saturation of the surface with adsorbed molecules.

The mechanism suggested for the reaction is that molecules from the gas phase react with molecules adsorbed on the surface, when these latter molecules acquire the necessary energy of activation from the metal atoms.

A remarkable and unusual similarity is found between the different metals in respect of (*a*) relation between reaction rate and pressure, (*b*) heat of activation, and (*c*) absolute rate of reaction. This would have suggested that the reaction really occurred in a zone of heated gas, were it not for the evidences that

surface saturation controls the rate of reaction. The aldehyde adsorption is evidently of a loose, non-specific kind. The hot wire does not appreciably modify the stability of the adsorbed aldehyde molecules (the heat of activation, being nearly the same as for the homogeneous reaction), but merely acts as a source of energy. Examination of the molecular statistics of the reaction indicates that collisions between molecules from the gas and adsorbed molecules are inelastic, with a duration of 10^{-6} to 10^{-8} second.

A Quantitative Study of the Reflexion of X-Rays by Sylvine.

By R. W. James, M.A., Senior Lecturer in Physics, and G. W. Brindley, B.Sc., Manchester University.

(Communicated by W. L. Bragg, F R S —Received July 26, 1928)

Introduction.

1. In recent papers[*] on the intensity of reflexion of X-rays from rocksalt crystals it has been shown that, from the temperature of liquid air up to about 500° abs., the dependence of the intensity of reflexion upon temperature is in accord quantitatively with a formula of the type originally deduced by Debye,[†] if the modification suggested later by Waller[‡] is introduced, although the decrease of intensity for higher temperatures is much greater than that indicated by the law. In dealing with quantitative results of experiments on reflexion from crystals, it is convenient to consider the quantity usually denoted by F, which is a measure of the scattering power, in a given direction, of an atom for X-rays. In the course of experiments with rocksalt it has been possible to determine F both for Na and Cl, and the values so obtained, when corrected for temperature, agree very closely with the F factors calculated from Hartree's Schrodinger density-distribution for the ions Cl^- and Na^+. The calculation is based upon the theoretical result, due to Wentzel[§] and Waller[||] that, to obtain the coherent scattering from an electron in an atom, the electron

---

[*] James, 'Phil. Mag,' vol. 49, p 585 (1925); James and Firth, 'Roy. Soc. Proc.,' A, vol. 117, p. 62 (1927); Waller and James, 'Roy. Soc. Proc.,' A, vol. 117, p. 214 (1927).

[†] 'Ann. d. Physik,' vol. 43, p. 49 (1914).

[‡] Upsala Dissertation, 1925.

[§] 'Z. f. Physik,' vol. 43, pp. 1, 779 (1927).

[||] 'Phil. Mag.,' vol. 4, p. 1228 (1927).

must be represented by its corresponding Schrodinger charge-density, each element of which must be supposed to scatter classically. In order to get agreement between the calculated and observed F curves, it is necessary to assume the existence of zero-point energy* of amount half a quantum per degree of freedom, which is of course required by the new quantum mechanics.

The agreement between theory and experiment in the case of rocksalt is extremely interesting, but, in order to place the quantitative treatment of X-ray reflexion on an entirely satisfactory basis, it appears to be of some importance to see whether a similar agreement can be obtained with other crystals. This is not quite so easy as might be supposed. If accurate absolute measurements of intensity are to be obtained from crystals, a number of conditions must be fulfilled which rarely all hold good in the same crystal. It is desirable that the positions of the atoms should be fixed by symmetry, so that there are no uncertain parameters to affect the intensity; the crystal must approximate to the ideal mosaic type,† so that the uncertainty due to primary and secondary extinction is not too great; moreover, the absorption coefficient for the X-rays used must not be too large, otherwise errors due to any non-crystalline layer on the surface, produced by grinding or by weathering, or to surface irregularities of any kind, will be too great. Some of these difficulties are to be avoided by using the crystal in the form of a powder, but it does not at present appear possible to obtain entirely satisfactory quantitative measurements from powdered crystals, although much valuable work on these lines has been done by Havighurst and others. In the work to be described in this paper we have used sylvine, KCl, which, although not as satisfactory as rocksalt, does fulfil a number of the necessary conditions, and in addition possesses the advantage that the two atoms are nearly of the same size and scattering power, so that the lattice approximates to a simple cubic one to which the theory of the effect of heat motion upon the intensity of reflexion, developed by Debye and Waller, properly applies. KCl possess the disadvantage that grinding or cutting faces appears to affect the surface to a certain extent. Although the intensity of reflexion from a freshly cleaved face is in general increased considerably by grinding, as in the case of rocksalt, there is some evidence that too rough a treatment of the face prevents the intensity from rising as high as it should. By cutting or grinding the necessary faces, and afterwards etching away the surface layer with water, it was found possible to get fairly consistent results. There remains the possibility that the intensities are all reduced by the presence of a non-crystalline surface layer, but the error due to this cause should not amount to more than a few per cent.

---

* James, Waller and Hartree, 'Roy. Soc Proc,' A, vol 118, p. 334 (1928).

† *Cf.* Bragg, Darwin and James, 'Phil. Mag.,' vol. 1, p. 897 (1926).

2. *Experimental.*

The method used in these experiments is exactly the same as that described previously for the case of rocksalt. For the high temperatures the crystal was mounted on a cast-iron holder inside a small furnace having mica windows through which the incident and reflected X-ray beams pass without appreciable absorption. The furnace is supported on a non-conducting stand which rests on the table of the X-ray ionisation spectrometer, and can easily be removed and replaced by the standard rocksalt crystal, with the (400) spectrum of which all the reflexions are compared. For the method of making the comparison and of allowing for the general radiation reference may be made to the papers already quoted. For the work at low temperatures, the crystal was supported in a brass holder just above the surface of liquid air in a thin-walled Dewar flask, the method of procedure being that developed by one of us with Miss Firth (*loc. cit.*). Observations were made at the temperature of liquid air and at a series of temperatures from that of the laboratory up to about 900° abs. The temperatures were measured with a calibrated iron-constantan thermo-couple.

In the work with rocksalt at high temperatures some difficulty was encountered owing to the fact that heating the crystal reduced the intensities of some of the spectra permanently. The effect was most marked for the strong spectra and was scarcely appreciable for the weaker ones, and appeared to be due to an internal recrystallisation, which increased the extinction of the crystal. Experiments were carried out with sylvine to see whether the same thing happened. No effect of the kind was observed, the intensity of reflexion from a given face remaining unaltered after several alternate heatings and coolings.

3. *Results of Experiments and Comparison with Theory.*

Observations both at high and low temperatures were made using the spectra (400), (600), and (800). The results for these faces are given in Table I. All the observations were made with MoK_α radiation ($\lambda = 0{\cdot}710$ Å.)

Table I.

Spectrum.	$T_{abs.}$	(ρ_T/ρ_{290}).	$(\rho_T/\rho_{290})_{corr}$	$\left(\frac{\lambda}{\sin\theta}\right)^2 \log_{10}\frac{\rho_T}{\rho_{290}}$
(400) $\sin\theta = 0·2270$	86	1·22	1·24	0·918
	290	1·00	1·00	0·000
	455	0·81	0·79	−1·006
	543	0·74	0·72	−1·403
	663	0·60	0·58	−2·326
	807	0·45	0·43	−3·602
	936	0·36	0·336	−4·656
(600) $\sin\theta = 0·3407$	86	1·607	1·624	0·9216
	412	0·73	0·73	−0·5982
	588	0·44	0·44	−1·560
	652	0·35	0·35	−1·995
(800) $\sin\theta = 0·4543$	86	2·46	2·46	0·9600
	336	0·78	0·78	−0·2650
	430	0·48	0·48	−0·7830
	471	0·38	0·38	−1·032
	587	0·21	0·21	−1·665

In this table, the third column gives the observed ratio of the intensity at a temperature T to that at the standard temperature, 290° abs. In the fourth column is given the same ratio corrected where necessary for the change of angle of reflexion with temperature, as described in an earlier paper,[*] and also for extinction in the manner described in section 7. We shall first compare these results with those to be expected from theory.

KCl approximates to the "mosaic" type of crystal and, for such a crystal, the intensity of reflexion, which is conveniently measured by the integrated reflexion, ρ, is given by the formula

$$\rho = \frac{N^2\lambda^3}{2\mu}\,\frac{\varepsilon^4}{m^2c^4}\,(\Sigma\,\mathrm{F}e^{-\mathrm{M}})^2\,\frac{1+\cos^2 2\theta}{2\sin 2\theta} \tag{1}$$

In this formula, ρ is the ratio of the energy received by the ionisation chamber, when the crystal is rotated through the reflecting angle, to that received in one second when the direct beam is allowed to enter the chamber. N is the number of crystal units per unit volume of the crystal, ε and m are the charge and mass of the electron, and c is the velocity of light. θ is the glancing angle of reflexion of the rays, so that 2θ is the angle of scattering, μ is the linear absorption coefficient of the crystal. If the crystal is "imperfect," but still shows secondary extinction, μ must be taken as the ordinary absorption coefficient increased by an amount which, for a given specimen of crystal, depends upon the strength of the reflexion.

---

[*] James, 'Phil. Mag.,' *loc. cit*

The factor F is the scattering factor for the atom, and for the purpose of this paper may be defined as the ratio of the amplitude of the coherent radiation scattered from an atom in a state of rest to the amplitude scattered by a free electron according to the classical theory due to J. J. Thomson. F is a function of $(\sin\theta)/\lambda$ and approaches a value equal to the total number of electrons in the scattering atom for small values of θ. The factor e^{-M} is the temperature factor, which allows for decrease in intensity of reflexion due to the heat motions of the atoms. The summation (ΣFe^{-M}) is to be taken over all the atoms in the crystal unit, and M will, in general, have a different value for each kind of atom. The crystal of KCl, in which the two kinds of atom have very nearly the same size and mass, can be treated as if its lattice were simple cubic with a single value of M. In this case we can apply the simple formula, deduced originally by Debye and modified by Waller, according to which

$$M = \frac{6h^2}{mk\Theta}\left\{\frac{\phi(x)}{x} + \frac{1}{4}\right\}\frac{\sin^2\theta}{\lambda^2}. \tag{2}$$

In this expression h is Planck's constant, k is Boltzmann's constant, m is the mass of each atom, which here is to be taken as the mean of the masses of K and Cl, and Θ is the "characteristic temperature." $x = \Theta/T$, and the function $\phi(x)$ is given by Debye in his paper, from which we have taken the values used here. The fraction $\frac{1}{4}$ added to $\phi(x)/x$ makes allowance for the existence of zero-point energy, but its presence or absence does not matter for the immediate purpose to which we shall apply this formula.

Neglecting small changes in θ due to the changes in dimension of the lattice with temperature, it follows at once from formulæ (1) and (2) that

$$\left(\frac{\lambda}{\sin\theta}\right)^2 \log\left(\frac{\rho_2}{\rho_1}\right) = \frac{6h^2}{mk\Theta}\left\{\frac{\phi(x_1)}{x_1} - \frac{\phi(x_2)}{x_2}\right\}, \tag{3}$$

where the suffixes 1 and 2 refer to two different temperatures. Putting $\Theta = 230°$, $m = 6 \cdot 149 \times 10^{-23}$ gm., the mean mass of the atom of K and Cl, and using the values of $\phi(x)$ given by Debye, we have calculated the values of $(\lambda/\sin\theta)^2 \log(\rho_T/\rho_{290})$ for a series of values of T from formula (3).

The dotted curve in fig. 1 is drawn from these values, and we have also plotted the observed points obtained from the last column of Table I. Fig. 1 may be compared with the similar figure aready given for rock salt.* It will be seen that, as in that case, the agreement between theory and experiment is very good up to about 400° abs. Above that temperature, the decrease of intensity

---

* James and Firth, *loc. cit.*

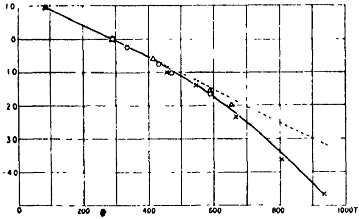


FIG. 1.—Ordinates represent $\lambda/\sin^2\theta \ \log_{10}(\rho_T/\rho_{290}) \times 10^{16}$.

with temperature is much more rapid than the formula indicates. It is not surprising that the formula does not hold good at high temperatures. In deducing it, the assumption is made that powers of the displacements greater than the second can be neglected in the expression for the energy. This can hardly be the case at high temperatures, when the amplitudes of the atomic displacements become very considerable.*

In the case of rocksalt, it was found that the experimental results at high temperatures were consistent with an empirical temperature factor e^{-A}, where

$$A = 1{\cdot}162 \times 10^{-5} \sin^2\theta \, . \, T^2/\lambda^2,$$

λ being expressed in Angstrom units. No particular reason can be given for the occurrence of T^2 in the exponent, but it is perhaps worthy of note that for KCl the results at temperatures above 400° abs. can be expressed by an empirical law of exactly the same type, the numerical factor in this case being $1{\cdot}330 \times 10^{-5}$.

In making a more detailed comparison of the experimental results with theory we shall confine ourselves to the observations made at room temperature and at the temperature of liquid air, since, over this range, a formula of the theoretical type holds good.

4. *Detailed Discussion of Results at Low Temperatures.*

The results of the measurements at the temperature of liquid air are summarised in Table II, in which the third and fourth columns give the values of

---

* *Cf* Waller, 'Ann. d. Physik,' vol. 83, p. 154 (1927).

Table II.

Spectrum	$h^2+k^2+l^2$.	(ρ_{86}/ρ_{290}).	$(\rho_{86}/\rho_{290})_{\text{corr}}$	ΔM.	$\Delta M/(h^2+k^2+l^2)$.
400	16	1·22	1·24	0·107	0·00672
600	36	1·61	1·62	0·242	0·00673
444	48	1·97	1·99	0·344	0·00717
800	64	2·46	2·46	0·450	0·00703
10,00	100	4·17	4·17	0·714	0·00714
666	108	4·82	4·82	0·786	0·00728
				Mean	0·00701

the ratio ρ_{86}/ρ_{290} for a series of spectra, as observed, and when corrected for extinction. From these ratios ΔM, the difference in the value of M at the two temperatures, can be calculated at once, for, from equation (1), it follows that

$$\rho_{290}/\rho_{86} = e^{-2\Delta M},$$

if we neglect the changes in the quantities which are functions of θ, produced by the small changes in the glancing angle due to the change in dimensions of the lattice with temperature. The numbers in the last column of Table II give the values of ΔM obtained from the observed ratios divided by $(h^2+k^2+l^2)$, (h, k, l) being the indices of the corresponding spectrum.

If a is the side of the full unit cube of the lattice

$$(\sin^2\theta)/\lambda^2 = (h^2+k^2+l^2)/4a^2,$$

and we obtain from equation (2),

$$\frac{\Delta M}{h^2+k^2+l^2} = \frac{3h^2}{2a^2 mk\,\Theta}\left\{\frac{\phi(x_1)}{x_1} - \frac{\phi(x_2)}{x_2}\right\}, \qquad (5)^*$$

where the suffixes 1 and 2 refer to two temperatures T_1 and T_2 for which the difference of M is ΔM.

From equation (5) it is seen that, for a given pair of temperatures, $\Delta M/(h^2+k^2+l^2)$ should be constant for all spectra The values shown in the last column of Table II are nearly constant, the differences in fact are practically within the limits of experimental error, and qualitatively theory and experiment are in accord. To make a quantitative comparison we have taken the values of $\phi(x)/x$ from the table given by Debye (*loc. cit.*), whence $\phi(x)/x = 1\cdot033$ and $0\cdot194$ at 290° and 86° respectively. Now for KCl, m = mean mass of K and Cl, $= 6\cdot149 \times 10^{-23}$ gm., $a = 6\cdot276 \times 10^{-8}$ cm., and Θ, the characteristic temperature, is 230°. Substituting these values in equation (5), and using the accepted values of h and k we obtain

---

* h and k each have two significances in equation (5), but confusion can hardly arise, and it seems inadvisable to alter a standard notation.

$$\Delta M/(h^2 + k^2 + l^2) = 0 \cdot 00706.$$

This value may be compared with the mean experimental value 0·00701. The agreement is very close, and is probably partly accidental, but it is satisfactory to find that, with a crystal such as KCl, which can justifiably be treated as simple cubic, the agreement with a formula of the type involving the characteristic temperature is closer than in the case of rocksalt, where the atoms differ considerably both in size and mass, so that a formula of this type should not apply so exactly.

5. *The Relationship between the Temperature Coefficients and the Elastic Constants.*

It is possible to compare theory and experiment using, instead of the characteristic temperature, the elastic constant. A comparison of this type has already been made for the case of rocksalt by Waller and one of the writers (R. W. J.), and the methods to be used in this section follow very closely those already described in their paper, to which reference may be made for details.

We may write

$$M = 8\pi^2 \sin^2\theta \overline{u_x^2}/\lambda^2, \qquad (6)$$

where $\overline{u_x^2}$ is the mean of the squares of the displacements, in any one direction x, of the atoms from their mean positions. For a cubic lattice $\overline{u_x^2}$ is independent of the direction of x. If the temperature T is high enough ($T > \Theta/2\pi$) we may write

$$\overline{u_x^2} = \alpha + \beta T + \gamma/T + \delta/T^3. \qquad (7)$$

In this formula, $\gamma = (h/2\pi)^2/12mk$, while δ can be evaluated approximately from the atomic constants, its value for KCl being with sufficient accuracy $-0 \cdot 56 \times 10^{-14}$. If we assume the lattice to possess vibrational energy at the absolute zero of temperature of amount half a quantum per degree of freedom, the term α is equal to zero. For a simple cubic lattice consisting of atoms of only one kind having mass m the following expression can be given for β,

$$\beta = \frac{k}{3m}\,\frac{\sqrt[3]{6\pi^2}}{2\pi^2} \cdot \rho a_0^{\,2}\,\frac{c_{44}(2c_{11} + c_{44}) + \frac{1}{3}b_1(c_{11} + c_{12})}{c_{11}c_{44}^{\,2} + \frac{1}{3}b_1(c_{11} + c_{12})\,c_{44} + \frac{1}{108}b_1^{\,2}b_2}, \qquad (8)$$

in which $b_1 = c_{11} - c_{12} - 2c_{44}$, $b_2 = c_{11} + 2c_{12} + c_{44}$, c_{11}, c_{12} and c_{44} being the elastic constants for the crystal in the notation of Voigt, ρ is the density of the crystal, and a_0 the edge of the elementary cube.[*]

Since the two kinds of atom in sylvine are so nearly equal in diffracting power, the spectra in which K and Cl are acting in opposition to one another are too small to be observed. We cannot therefore determine M for the two atoms separately, as was possible with rocksalt; but since the atoms are also very nearly equal in mass and size we can treat the lattice as simple cubic and apply formula (8) putting $m = \frac{1}{2}(m_{\text{K}} + m_{\text{Cl}})$. Since the cube of edge a_0 contains half a molecule of KCl we can simplify (8) by putting

$$\rho a_0^{\,3} = \tfrac{1}{2}(m_{\text{K}} + m_{\text{Cl}}) = m.$$

For KCl, $a_0 = 3\cdot136 \times 10^{-8}$ cm., and, from the work of Voigt,[†] we find $c_{11} = 3\cdot680 \times 10^{11}$, $c_{12} = 0\cdot194 \times 10^{11}$, $c_{44} = 0\cdot683 \times 10^{11}$. The value of γ for KCl is $1\cdot075 \times 10^{-17}$. Using these values, we find from equation (8)

$$\boldsymbol{\beta} = \mathbf{6\cdot96 \times 10^{-21}}.$$

The calculated value of β may now be compared with that which can be deduced from the experimental results given in Table II. From (6) and (7), expressing $(\sin\theta)/\lambda$ in terms of the indices (h, k, l) of the spectrum, and the edge a of the unit cell, we have

$$\frac{\Delta\text{M}}{h^2+k^2+l^2} = \frac{2\pi^2}{a^2}\left\{\beta\,(\text{T}_1 - \text{T}_2) + \gamma\left(\frac{1}{\text{T}_1} - \frac{1}{\text{T}_2}\right) + \delta\left(\frac{1}{\text{T}_1^{\,3}} - \frac{1}{\text{T}_2^{\,3}}\right)\ldots\right\}. \quad (9)$$

Substituting in this equation the experimental value of $\Delta\text{M}/(h^2+k^2+l^2)$ from Table II, and the calculated values of γ and δ, we obtain at once

$$\boldsymbol{\beta} = \mathbf{7\cdot25 \times 10^{-21}}.$$

The calculation using the elastic constants gave

$$\boldsymbol{\beta} = \mathbf{6\cdot96 \times 10^{-21}}.$$

The agreement is again quite as close as can be expected, and the results obtained with rocksalt are entirely confirmed by this investigation.

On the assumption that the lattice possesses zero-point energy, so that $\alpha = 0$ in equation (5), we may obtain from the experimental results,

$$\text{M}/(h^2+k^2+l^2) = \frac{2\pi^2}{a^2}\left\{7\cdot25 \times 10^{-5}\,\text{T} + \frac{0\cdot1075}{\text{T}} - \frac{56}{\text{T}^3}\right\}, \qquad (10)$$

---

[*] This formula is due to Dr. I. Waller, to whom we are much indebted for permission to use it in this paper prior to its publication elsewhere.

[†] 'Ann. d. Physik,' vol. 36, p. 743 (1889).

in which a is expressed in Angstrom units.* From this formula, M can be calculated for any temperature.

Still assuming the existence of zero-point energy, we may now calculate from equations (6) and (7) the values of the mean square amplitude of vibration of the atoms of the lattice at any temperature.

At 290° abs. we have

$$\overline{u_x^2} = 0 \cdot 0214 \text{ Å}^2$$

and at 86° abs.

$$\overline{u_x^2} = 0 \cdot 0074 \text{ Å}^2.$$

If now $\overline{u^2}$ is the mean of the squares of the total displacements of an atom, for a cubic crystal

$$\overline{u^2} = \overline{3u_x^2}.$$

Hence for KCl at 290°

$$\sqrt{\overline{u^2}} = 0 \cdot 255 \text{ Å},$$

and at 86°

$$\sqrt{\overline{u^2}} = 0 \cdot 149 \text{ Å}.$$

6 *Comparison of Observed Results with Calculations from Atomic Models.*

We must now investigate whether or not the experimental results indicate the existence of zero-point energy. In order to do this it is necessary, as in the work on rocksalt, to determine the absolute values of the scattering coefficients for the atoms of K and Cl by experiment, and compare them with the values calculated from an atomic model. With sylvine only spectra of the type K + Cl can be observed. The intensities of reflexion were measured in the usual way, the (400) reflexion from rocksalt being taken as the standard. Since the absolute value of the reflexion from this face is known the absolute value of any other spectrum, compared with it, can be determined. From the observed intensities, the values of $(\text{F}e^{-\text{M}})_{\text{K}} + (\text{F}e^{-\text{M}})_{\text{Cl}}$ are calculated using formula (1). Observations were made at room temperature and at the temperature of liquid air. The spectra of very high order were observed at the low temperature only. The apparatus was modified to allow spectra to be observed at angles of scattering as high as 130°, and, using rhodium radiation, it was possible to obtain the spectrum (18,00), although no reliable measurement of its intensity could be made.

In columns (2) and (3) of Table III the absolute values of the integrated reflexions, ρ, for a series of spectra at 290° and at 86° abs. are given. The values of $(Fe^{-M})_K + (Fe^{-M})_{Cl}$, calculated from these intensities by means of equation (1), are given in the last four columns those in columns (4) and (5) are the uncorrected values, calculated using the ordinary absorption coefficient, while those in columns (6) and (7) have been corrected for extinction.

---

* In the corresponding formulæ (6) and (6A) of the paper by Waller and James, the second term on the right-hand side should in each case contain the factor 10^{-1} instead of 10^{-3}. The correct value was used in all the calculations made in that paper.

Table III.

Spectrum.	$\rho \times 10^6$.*		$\Sigma (Fe^{-M})$			
			Uncorrected.		Corrected for extinction	
	290° abs.	86° abs	290° abs.	86° abs	290° abs	86° abs
200	275	—	21·4	-	26·1	(26·8)
222	93·7	—	16·8	—	17·8	(19·3)
400	60·0	73·4	14·5	16·2	15·1	17·0
600	14·6	23·5	9·22	11·7	9·31	11·9
444	8·41	16·6	7·72	10·9	7·76	11·0
800	4·26	10·5	6·08	9·60	6·10	9·65
10,00	1·10	4·60	3·64	7·45	3·64	7·48
666	0·81	3·90	3·19	7·04	3·19	7·06
12,00	—	1·83	—	5·07	—	5·07
888	—	1·04	—	3·71	—	3·71
14,00	—	1·10		3·76	—	3·76
16,00	—	0·64	—	2·18	—	2·18

* The absolute value of ρ_{400} NaCl, which was used as the standard in these observations, was taken as 98·4 × 10^{-6}.

To compare these results with theory, we have calculated the Schrodinger charge-density distribution for the potassium ion, K^+, using the method developed by Dr. Hartree,* to whom we are greatly indebted, for supplying us with the initial values of the field for starting the calculations, and for much help and instruction during its actual course. By supposing each element of this density distribution to scatter classically, the F curve for the K^+ ion in a state of rest can be calculated. For details of the calculation of F by this method reference may be made to the paper on rocksalt.

In fig. 2, the charge distribution for K^+ is plotted, together with that already obtained for Cl^- for comparison. The ordinates of the curve give the values of U (*r*), where U (*r*) *dr* is the total charge, expressed in terms of the electronic charge as unit, between spheres of radii *r* and *r* + *dr*. The abscissæ give the distance from the centre of the atom in Ångstrom units. Each ion contains 18 electrons; the areas beneath the two curves are therefore equal, but the greater nuclear charge of K^+ causes the scale of the distribution to shrink.

---

* 'Camb. Phil. Soc. Proc.,' vol. 24, pp. 89, 111 (1928).

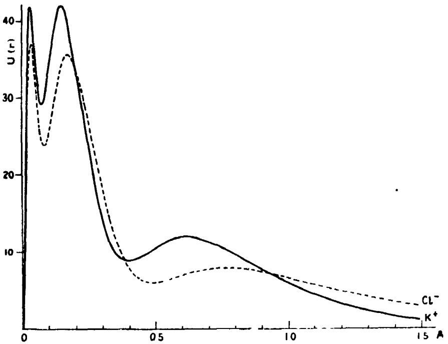


Fig. 2.—Schrödinger Charge-Density Distributions for K^+ and Cl^-.

The two F curves resulting from these distributions are shown in fig. 3, and the numerical values are given in Table IV. Each tends to the same limit,

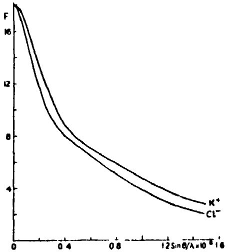


Fig. 3.—F Curves for Ions of K^+ and Cl^-.

18, for $(\sin\theta)/\lambda = 0$, but the greater compactness of the K^+ ion causes its F curve to fall away more slowly as $(\sin\theta)/\lambda$ increases. The two F curves are,

Table IV.—Calculated Values of F for the Ions K^+ and Cl^- at Rest.

$\frac{\sin\theta}{\lambda}\times 10^{-8}$.	K^+	Cl^-	$\frac{\sin\theta}{\lambda}\times 10^{-8}$.	K^+	Cl^-
0·00	18·00	18·00	0·50	7·77	7·23
0·05	17·72	17·11	0·60	7·05	6·49
0·10	16·46	15·23	0·70	6·44	5·77
0·15	14·82	13·19	0·80	5·90	5·06
0·20	13·31	11·50	0·90	5·32	4·41
0·25	12·00	10·23	1·00	4·70	3·84
0·30	10·78	9·30	1·10	4·22	3·33
0·35	9·78	8·60	1·20	3·73	2·89
0·40	8·83	8·06	1·30	3·32	2·51
0·45	8·28	7·62	1·40	2·98	2·21

however, nearly enough equal to account for the absence of any certain spectra of the type K—Cl. The intensity of a spectrum of this type at any angle will be a fraction $(F_K - F_{Cl})^2/(F_K + F_{Cl})^2$ of the normal value at that angle. This fraction has very small values for small values of $(\sin\theta)/\lambda$. For larger angles of scattering the value of the fraction is considerable, but here the normal value is so small that the spectrum would still be too weak to be observed.

From the F curves for K^+ and Cl^- the values of $F_K + F_{Cl}$ for the atoms in a state of rest can be obtained, and hence, using the value of M already determined, the value of $(F_K + F_{Cl})e^{-M}$ can be calculated for any temperature. If the value of M given by equation (8) is used, this assumes the existence of zero-point energy. An approximate value of M assuming there to be no zero-point energy can be obtained from formula (2) by omitting the term $\frac{1}{2}$.

In Table V are given the values of $(F_K + F_{Cl})\,e^{-M}$ for a series of values of $\sin\theta/\lambda$, calculated from the model atom, and reduced to the temperature of liquid air. The values in column (4) assume the existence of zero-point energy, those in column (3) its absence. In column (5) are tabulated the observed values. If we confine our attention for the moment to large values of $\sin\theta/\lambda$, it will be seen that the observed values agree very well with those calculated assuming the existence of zero-point energy. As was pointed out in the paper on rocksalt, it is the value of F for large values of $\sin\theta/\lambda$ which are of importance in deciding for or against zero-point energy, for these values depend almost entirely on the K electrons of the atoms, whose Schrodinger density can be calculated with some approach to certainty. Moreover, the amplitude of vibration corresponding to zero-point energy is considerably greater than the radius of the region within which the main part of the density distribution representing the electrons will be confined when the atoms are at rest. The

Table V.—Values of $(F_K + F_{Cl})e^{-M}$ at 86° abs.

Spectrum.	$\frac{\sin\theta}{\lambda} \times 10^{-8}$.	Calculated.		Observed.
		No zero-point energy.	With zero-point energy.	
200	0·160	27·08	26·81	(26·8)
222	0·275	20·44	20·09	(19·3)
400	0·319	18·99	18·36	17·0
600	0·478	14·50	13·45	11·9
444	0·555	13·16	11·85	11·0
800	0·638	11·74	10·27	9·65
10,00	0·797	9·35	7·58	7·48
666	0·834	8·85	7·04	7·06
12,00	0·957	7·21	5·33	5·07
888	1·113	5·46	3·64	3·71
14,00	1·116	5·39	3·58	3·76
16,00	1·276	3·98	2·33	2·18
18,00	1·435	2·96	1·38	—

difference between the calculated F curve for an atom at rest, and for one vibrating with this amplitude, is therefore very considerable.

7. *The Correction for Extinction.*

For the comparison between theory and experiment at large angles of scattering no correction for extinction has been made. For the smaller angles, at which the intensities of the spectra are much greater, some correction is necessary. Sylvine like all crystals is neither ideally perfect nor ideally imperfect (or "mosaic").* It approximates fairly closely to the "mosaic" type, but there is certainly some secondary extinction, and possibly some primary. The extinction has not been measured for this crystal, and to get an approximate correction for its effects we have proceeded in the following manner: From the calculated F curves for K^+ and Cl^-, corrected for temperature, the value ΣFe^{-M} for the (200) spectrum at the temperature of the laboratory may be calculated. It is found to be 26·07. The measured value at the same temperature is 21·4. We have supposed the defect to be entirely due to secondary extinction. There is some warrant for doing this, since the calculated and observed values are in agreement for large angles of scattering, that is to say, for the weak spectra, for which the secondary extinction is in any case negligible. According to Darwin,† we may allow for secondary extinction if we write formula (1) in the form

$$\rho = Q/2\mu,$$

---

* *Cf.* Bragg, Darwin and James, *loc. cit.*

† 'Phil. Mag.,' vol. 43, p. 800 (1922).

and take μ to be equal to $\mu_0 + gQ$, where μ_0 is the ordinary linear absorption coefficient for X-rays passing through the crystal, not at the reflecting angle, and g is a constant for the given specimen of crystal. The more "perfect" the crystal the larger g. This has the effect of increasing the absorption for the strong spectra. If we suppose that the difference between the observed and calculated F for (200) to be due to secondary extinction, we can make an estimate of g, for we shall have

$$\frac{\mu_0 + gQ}{\mu_0} = \left(\frac{26 \cdot 07}{21 \cdot 41}\right)^2.$$

Now Q is known for (200) and μ_0 can be measured. We have assumed a value $\mu_0 = 27 \cdot 0$, the mean of a large number of concordant determinations.

Hence g can be calculated, and the value of the extinction determined, for any spectrum whose intensity is known. The values of F corrected for extinction in this way are shown in columns (6) and (7) of Table III. It will be seen that the correction is very small, except for the strong spectra. This method of correcting for extinction, which was first used by one of the writers with J. T. Randall* for the case of fluorite, is, of course, not very satisfactory, in that it forces the observed and calculated curves to agree at one point. In the absence of any measured data for the extinction of KCl, it appears to be the best that can be done, and it should be emphasised that the agreement between theory and experiment for large angles of scattering is independent of any such correction. The correction for extinction introduced in determining the temperature factor may be explained here. Since the intensities differ at the two temperatures, the effective absorption coefficients differ also, and it is just worth while to apply a correction for this effect. This has accordingly been done in column (4) of Tables I and II.

In fig. 4 we have plotted the calculated curve (A) for $F_K + F_{Cl}$ when the atoms are at rest together with those for $(F_K + F_{Cl})e^{-M}$ for the temperature of liquid air. Curve B assumes the existence of zero-point energy, curve C its absence. The circles show the observed values. There are two points of interest. The first is that, for high angles of scattering, there is no doubt that the experimental and theoretical results agree very well, assuming the presence of zero-point energy. As the new quantum mechanics, upon which the calculation of the F curves is based, requires this energy it is satisfactory to find this agreement. The result is also in accord with that obtained for rocksalt. The second point is that for smaller values of $(\sin \theta)/\lambda$ the observed points lie markedly lower than the calculated curve. This was found to be the case both for sodium and for chlorine, though to a greater degree for chlorine than for sodium. For KCl the effect is even more pronounced, and here of course both ions have the same set of "orbits." In the region of (sin θ)/λ where the deviation occurs the outer electrons still have a large influence upon the value of F; it is perhaps not basing too much upon the experimental results to suggest that the method used to obtain the field acting on these electrons may need some modification.

---

* James and Randall, 'Phil. Mag.,' vol. 1, p. 1202 (1926).

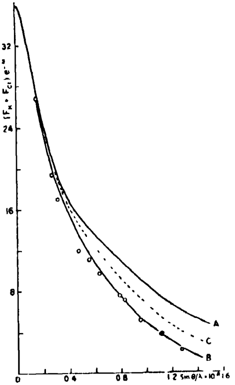


FIG. 4

Summary.

This paper is an extension to the case of sylvine of quantitative work already published for the case of rocksalt.

1. The variation of the intensity of reflexion of X-rays from sylvine with temperature has been shown to be quantitatively in agreement with the Debye-Waller law from the temperature of liquid air up to about 400° abs. At

higher temperatures, the decrease of intensity with increasing temperature is much more rapid than the law indicates.

2. The value of the temperature-factor based on observations at room-temperature and at the temperature of liquid air agrees very well with that calculated by Waller from the elastic constants of the crystal, and also with the value calculated from the Debye-Waller law using the characteristic temperature.

3. The absolute values of the intensity of reflexion are in good agreement with those calculated from the Schrodinger density distribution for K^+ and Cl^- obtained by Hartree's method, if each element of the distribution is assumed to scatter classically, and if, in correcting for temperature, the existence of zero-point energy is assumed.

In conclusion we wish to express our thanks to Prof. W. L. Bragg, F.R.S., for his constant interest and advice during the course of this work. We are also indebted to Mr. C. H. Bosanquet, M.A., and to Mr J. M. Bruckshaw, B.Sc., for assistance with some of the observations, and to Dr. D. R. Hartree for much help in the calculation of the Schrodinger distribution. The sylvine used in these experiments was very kindly sent to us by Prof. A. Šimek of the University of Brno.

The Mobility of Ions in Air. Part IV.—*Investigations by Two New Methods.*

By A. M. Tyndall, L. H. Starr and C. F. Powell.

(Communicated by A. P. Chattock, F.R S.—Received July 26, 1928)

Introductory.

In a previous communication to the Royal Society (vol. 110, p. 342 (1926)) an account was given of measurements of mobility in air containing water vapour. The essential feature of the method used was the production of a layer of ions by a momentary flash of α-rays and their subjection to a special type of alternating field. When all the ions of one sign are of the same mobility the current is confined to a narrow range of frequency of alternation of the field, rising to a sharp maximum or " peak " at a frequency which is a measure of the mobility. The apparatus was used to investigate the effect of water and also of a series of organic vapours* on the mobility of air ions, and, incidentally, gave further evidence of the change which Erikson discovered in the nature of the positive ions with age.

In considering the extension of the experiments to other problems, certain disadvantages of the method were apparent. Firstly, to obtain the necessary flashes of α-rays in synchronism with the alternating field, the polonium source was mounted outside the ionisation chamber, and the α-rays admitted through a thin mica window. The apparatus, therefore, could not be made thoroughly airtight, and was not suitable for extremely dry air or for other gases. Secondly, the age of the ions was entirely governed by the time taken by them in crossing the mobility box. Therefore, not only was the age strictly limited, but also in the case of positive ions there was definite evidence of a transformation in the nature of some of the ions during the measurement itself.

Since the publication of this paper two other methods of mobility measurement have been devised and tested in this laboratory. A description of these methods and of the results obtained with them is the subject of the present paper.

---

* ' Roy. Soc. Proc.,' A, vol. 111, p. 577 (1926).

The Four Gauze Method

The first method will be referred to as the "Four Gauze" method.* Its main advantages are that the apparatus can be made airtight, that there is no commutator, that the method has a higher resolving power than the older method, that the time of measurement is shorter, and that the ions may be given any required age (within certain limits) before their mobility is measured. The method is not absolute, but for the purpose in mind only relative values of mobility were required.

Fig. 1 is a diagram of the apparatus The electrodes A to H are plane metal sheets carefully adjusted for parallelism. The central portions of B and C are coarse wire gauzes, and those of D, E, F and G are of perforated metal.

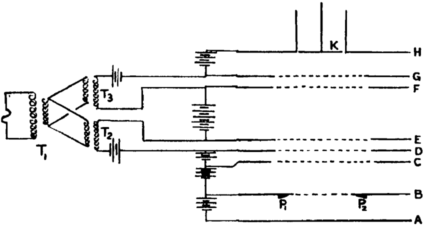


Fig. 1.

P_1 and P_2 are two strips of copper coated with polonium generating ions between A and B in a field which drives ions of the sign under investigation to the gauze B. Fields between B and C and between C and D carry some of these ions to the gauze D, and the age of the ions when they arrive at D may be given any required value (within certain limits) by suitably adjusting the strength of the field B C. The field C D is not essential to the method; it was inserted to prevent interpenetration between the ageing field B C and the alternating fields above D. Between E and D, which are close together, an alternating sinusoidal voltage is applied with an opposing direct voltage or

---

* While this communication was being prepared for publication, a paper by Mr. R. J. Van de Graaff appeared in the July number of the 'Philosophical Magazine' ('Phil. Mag.,' vol. 6, p. 210 (1928)), describing a method identical with this in principle. But our apparatus differs in several important respects from that of Mr. Van de Graaff, and on that account a detailed description of it is included.

"bias" superposed. As a result a narrow band of ions arrives at E when the advancing phase of the alternating field is near its peak voltage. This layer now enters a steady field applied between E and F, arriving at F after an interval of time, the value of which depends upon the ionic mobility and the values of the distance and voltage between E and F.

The electrodes F and G, and the voltages applied between them, are identical with E and D. Moreover, the two alternating fields are exactly in phase, being taken from the same generator through the system of transformers T_1, T_2 and T_3. Between G and H a steady potential difference takes any ions emerging from G on to an insulated electrode K, connected to the electrometer, or a guard ring H surrounding it. The result is that the ions will pass through F, G and on to the electrometer in every cycle, provided that the time they take to go from E to F is approximately equal to the duration of a cycle. Similarly they will reach the electrometer if the time of passage of the ions from E to F is that of two, three or more cycles, but their numbers will be considerably smaller. On the other hand, except in the immediate neighbourhood of these times, no ions will emerge. The procedure adopted is, therefore, to keep all the direct and root mean square alternating voltages constant, and to vary the frequency of alternation. This can be done by varying the speed of revolution of the generator and adjusting the current in its field coils. Under these conditions, if the ions are all of the same mobility, the electrometer current will be zero except in the neighbourhood of certain frequencies, and, if plotted with frequency, will show a rise to a maximum and a subsequent fall. Of these maxima, that corresponding to a time of passage of the ions from E to F in one cycle will be by far the most prominent. If, on the other hand, ions of different mobilities are present, each group should give its own series of "peaks" under ideal conditions of resolving power. The wide separation of the harmonic peaks from the fundamental enables the investigation of possible separate groups to be carried out over a fair range of mobility without interference.

For the investigation of ions undergoing rapid transformation after birth, it is essential that the time taken during the measurement of mobility itself should be reduced to a minimum. A limit was in practice placed upon this by the range of frequency of the alternator at our disposal. In most of the curves given in this paper the time of measurement was about 1/100 second. The values of the age of the ions given in this paper do not include the time of measurement.

All the measurements of mobility were made at atmospheric pressure.

Details of the Apparatus.

The apparatus was constructed in several forms, differing from each other only in dimensions and in the way in which the electrodes were supported. In the form of apparatus with which all the results given in this paper were obtained (except those for nitrogen and very dry air), the electrodes D to G consisted of brass plates, 14 cms. × 9 cms., in each of which there was a circular hole 7·5 cms in diameter. To each plate a sheet of perforated zinc was soldered so that it covered this hole. The electrodes A to H were supported on glass rods in a rectangular metal frame. The electrode K was of 3 cms. diameter, and was supported by a metal rod passing through a sulphur plug. The whole apparatus was enclosed in a bell jar on a glass plate.

The distances between the electrodes were as follows: A B = 1 cm. B C = 1 cm. C D = 0·7 cm. D E = 0·5 cm. E F = 2 cms. F G = 0·5 cm. C H = 1 cm. The alternating voltages D E and F G were of R M S 120 volts and the " bias " was usually 60 volts.

Results.

Negative Ions in Air.—The results for negative ions in air containing water vapour are given in Table I, in which values from Tyndall and Grindley's curve[*] are compared with those by the new method, assuming that the mobility in dry air is the same in both cases. The short and long ages used were 0·016 and 0·25 second respectively.

Table I.—Mobility of Negative Ions—Effect of Water Vapour.

Vapour pressure of water in mms. of mercury.	Tyndall and Grindley.	Short age.	Long age.
0·6	2·08	2·07	2·01
0·8	2·06	—	2·09
1·5	2·01	—	2·02
3·7	1·91	1·87	1·85
7·2	1·80	1·80	1·81
10·6	1·72	1·71	1·73

The results by the new method are not corrected to a standard atmospheric temperature and pressure, but the agreement is very satisfactory, and may be taken as showing the accuracy of the method. The results, moreover, indicate that the mobility of the negative ions is independent of age.

---

[*] ‘Roy. Soc. Proc.,’ A, vol. [illegible]10, pp. 9, 342 (1926).

In fig. 2 curve A is the ionisation-current-frequency curve from negative ions in air containing 10·6 mms. of water vapour, and is typical of all the negative peaks obtained. It shows the resolving power of the apparatus.

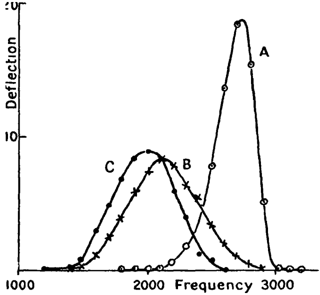


Fig. 2.—Ions in Air. Electrometer Deflection—Frequency (R.P.M.).
⊙ Short aged Negatives. × Short aged Positives. • Long aged Positives.

Positive Ions in Air.—The results for positive ions were quite unexpected. In fig. 2 curve B is for positive ions at an age of about 0·02 second in air containing 10·6 mms. of water vapour. From the results of Erikson and of Tyndall and Grindley, one would expect under these conditions a preponderance of initial ions, which would be shown by a maximum current at a frequency equal to that of the negative peak. In curve B the peak frequency is very nearly equal to that corresponding to final ions, and the curve suggests that although a few initial ions are present, the transformation into finals has occurred very much more rapidly than in Erikson's experiments. Broad curves indicating the presence of both initial and final ions were obtained at this age at a number of humidities (corresponding to water-vapour contents of from 0·6 to 10·6 mms.), and in nearly all cases the curves gave evidence of an unexpectedly high rate of transformation. In a great many cases (of which curve B is an example) the form of the curve suggested the presence not only of initial and final ions, but also of ions of intermediate mobility; but a quantitative study of these ions was rendered difficult by reason of the transformation taking place during the measurement.

Moreover, the results for positive ions, unlike those for negative, were variable. It was found on some occasions, when an experiment was repeated under the same conditions of age and humidity but after an interval of weeks, that the positive curve showed a different relative number of initial and final ions, while the negative curve was unchanged. Indeed, in one experiment an actual preponderance of initial ions of mobility equal to that of the negatives was observed under conditions of age and humidity in which, on all other occasions, there was a large preponderance of final ions. This variability in the rate of transformation of the positive ions was clearly due to some uncontrolled factor, and it was thought that this factor might possibly be the presence of small traces of some substance (such as a product of the chemical action of the alpha rays), the amount of which depended in some way on the state of the surfaces of the electrodes or of the walls. Further experiments bearing on this question are described in Part V.

While there is this variability in the results for short age positives in air containing water vapour, the results at long age are more definite. For positive ions of relatively long age, 0·25 second, the values of the peak frequency (which is proportional to the mobility) for a series of vapour pressures of added water are shown in Table II. The mobility values in the third column are calculated on the assumption that the mobility of the negative ion in dry air is 2·15.

Table II.

Mms. of water vapour.	Peak frequency.	Mobility.
0·6	1950	1 22
1 5	2000	1·25
3·7	1970	1·23
7·2	2000	1·25
10 6	2000	1 25

The results suggest that variation in water content over wide limits has no effect upon the mobility of the final positive ion. In fig. 2 curve C is for positive ions at an age of 0·25 second. This curve, which is similar to those obtained at other humidities, is broader relatively to the frequency at which it occurs than is the negative curve, and suggests that even at this age the positive ions are not a single group but are spread over a small range of mobilities. The values given in Table II refer, therefore, to the mean value of the mobility of the positive ions under these conditions.

Experiments in Nitrogen.—In view of the possibility referred to above that products of the chemical action of α-rays may play an important part in the transformation of positive ions, some experiments were carried out in pure nitrogen for the purpose of eliminating all complications which may arise from such products.

For this purpose the apparatus was constructed in a different form. The electrodes, which were of perforated copper cleaned in nitric acid, were separated by glass distance pieces, and special care was taken to secure their accurate parallelism and to align their perforations so that a maximum number of ions reach the electrometer. The apparatus rested on a sheet of plate glass and was covered by a bell jar. The electrical connections to the gauzes entered the apparatus through pinches in lead glass, carried by the sleeves of two conical ground glass joints passing through the base plate. Two glass tubes ground into two other conical holes in the plate formed the entrance and the exit tubes for the introduction and removal of gas. The electrometer lead passed through an amber plug ground into a conical hole in the top of the bell jar. All these joints were made airtight by vacuum wax applied externally.

The nitrogen, which was taken from a cylinder, was purified before it entered the apparatus by passing it over heated copper and through calcium chloride, phosphorous pentoxide and liquid air traps. Dishes containing phosphorous pentoxide were also placed under the bell jar. The apparatus was pumped out to a pressure of 10^{-5} mms. of mercury, and nitrogen was allowed to enter slowly until the pressure was 10 cms. These operations were repeated several times, and nitrogen was then allowed to enter the apparatus until the pressure was 720 mms. of mercury.

Some results are shown in fig. 3. The age of the ions was about 0·005 second, and the time of measurement about 0·005 second. Curve A is for negative ions taken immediately after the gas had been admitted. Curve B is the corresponding curve taken after 18 hours. Curve A shows that under these conditions about one-half of the ions present pass through the electrometer at all frequencies. In other words, they have a very high mobility and are, no doubt, electrons. The presence of electrons is in accordance with the work of Franck and others, and, incidentally, it provides a check on the purity of the gas. Some ordinary ions are, however, present, as is shown by the peak which occurs in about the same position as that given by air. Curve B shows a greatly reduced number of electrons, owing, no doubt, to the effect of small traces of impurity given off from the walls and electrodes in the time intervening between the two experiments. It is not unlikely, therefore, that in an apparatus which could be baked out these negative ions would disappear.

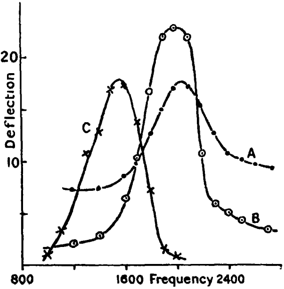


FIG. 3.—Ions in Nitrogen. Electrometer Deflection—Frequency.
● Negatives. ⊙ Negatives 18 hours later. × Positives.

Curve C is the corresponding curve for positive ions in fresh nitrogen. This curve did not change appreciably in 18 hours. The position of the peak corresponds to a mobility about equal to that of the final positive ions in air. It will be seen that there is no indication of the presence of any initial ions under these conditions. The bearing of these results on the nature of the transformation of the positive ions in air is discussed in Part V.

Experiments in Dry Air.—The apparatus in which the nitrogen results were obtained was also used to investigate the ions in very dry air. The air was dried in exactly the same way as the nitrogen.

The negative ions gave a sharp peak at a frequency corresponding to a mobility of about 2·15, and there was no indication of the presence of electrons. Now Lattey* working in dry air and Brose† in dry oxygen found high mobilities indicating a preponderance of electrons; but in the present apparatus, despite similar precautions to exclude all vapours, no indication of electrons was found. The apparent discrepancy may well be due to the fact that Lattey used pressures from 14 to 28 mms., and Brose still lower pressures. Under these conditions the concentration of products of chemical action due to the ionising agent must have been far less than that existing in the present apparatus.

---

* 'Roy. Soc. Proc.,' A, vol. 84, p. 173 (1910).

† 'Phil. Mag.,' vol. 50, p. 536 (1925).

The positive ions, at an age of 0·005 and a time of measurement of 0·005 second, gave only one peak, namely, that corresponding to the mobility of final ions. No initial ions were found under these conditions. This result is not surprising in view of the earlier experiments of Miss Valesek[*] and of Tyndall and Grindley (*loc. cit.*), in which it was found that the rate of transformation increases at low humidity, and of the fact that the present method gives a high rate of transformation even in moist air.

The Two Slit Method.

The second method was developed concurrently with the investigations described above in the hope of securing confirmation of the results obtained. This method also gives ions of controlled age and a peaked curve for ions of a single mobility. But it is not an absolute method, and it involves the use of a square wave field produced by a revolving commutator, so that the time of measurement is necessarily long. The following description, therefore, includes only a bare minimum of experimental detail.

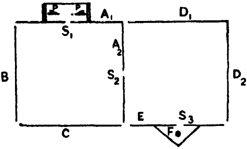


FIG. 4A.

Fig. 4A shows a section of the mobility chamber which was mounted on a glass plate and covered with a bell jar, the electrode connections passing through seals in the plate. $A_1 A_2$ is a metal plate bent at right angles, and containing two slits S_1 and S_2. B and C are metal plates completing with $A_1 A_2$ four sides of an open box. $D_1 D_2$ is a plate similar to $A_1 A_2$, but without slits. E is a plate with a slit S_3 giving entrance to an enclosed chamber containing a collecting rod F connected to a Compton electrometer. Ions from polonium at P are drawn continuously towards S_1 by steady fields. A steady field also operates between E and F. The fields between the various other electrodes are put on in a certain order by a commutator revolving at constant, known and controllable speeds. The form of the field may be regarded conveniently as passing through three phases shown in fig. 4B.

---

[*] 'Phys. Rev.,' vol. 29, p. 542 (1924).

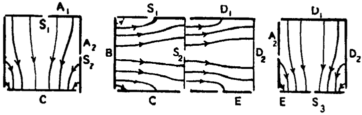


FIG 4B.

In Phase I, A_1, A_2 and B are at the same high potential V, and C is at half this potential. Ions from S_1 pass across to C in a thin layer, in which lateral diffusion is reduced by the convergence of the field. Phase I is then cut off by the commutator and Phase II brought on. B is then at the potential V; A_1, A_2 and C are at half this potential, and D_1, D_2 and E at a still lower potential. The ionic layer is thus driven to A_2, and a narrow band from it passes on through S_2 towards D_2. Again the commutator operates, and Phase III is brought on. In this phase A_2, D_1 and D_2 are at the full potential, and E at a low potential. If at the time the band of ions is anywhere between S_2 and D_2 it will be drawn towards E, and the actual position of the band when Phase III starts to operate will determine whether or not ions will pass through S_3 and reach the electrometer.

In practice, Phases I and III coincide in time so that a new layer is formed in one half of the box, while the remains of the previous layer are drawn to E in the other half.

The times for which these phases last are determined by the speed of the commutator. At low speeds of commutator the ions, during Phase II, go past the point at which Phase III would draw them to the slit S_3. Similarly at high speeds they do not reach this point. If the ions are all of the same mobility, and if the commutator speed is gradually changed, the electrometer current is zero except in a critical region, in which it rises to a peak value and falls again. As in the previous method, separate peaks should occur if groups differing in mobility are present.

The age of the ions could be varied from 1/65 to 2/3 second in the ionisation chamber before emission from S_1. The time occupied in travelling from S_1 to S_3 could not be reduced conveniently below 1/40 second, because it was found that at high speeds of revolution of the commutator, or in a commutator with short sectors, the time of contact of the brushes with the sectors was variable. Thus the range of age possible was considerably less than that of the four gauze method.

Results.

(*a*) *Air containing Water Vapour.*—In air with water vapour, the experiments gave a general confirmation of the results obtained by the four gauze method and in the previous method of Tyndall and Grindley.

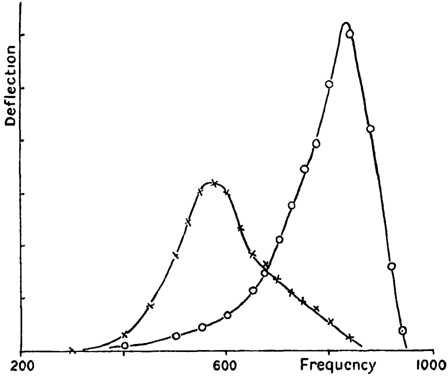


Fig. 5.—Two Slit Method. Ions in Air containing Water Vapour.
⊙ Negative. × Positive.

Fig. 5 gives as an illustration the results for air containing 3·6 mms. of water vapour. The resolving power of the apparatus is seen to be good and suitable for the detection of initial ions because of the form of the curve, which drops sharply on the high frequency side. Comparison of the curves for positives and negatives suggests the presence of some initial positive ions, with a general distribution in mobility similar to that of curve B in fig. 2, obtained by the four gauze method.

The method also confirmed the results at long ages in showing that the mobility of the final positive ion is independent of the amount of water vapour present. There seems, therefore, no reason to doubt the reliability of either method, and the necessity of seeking some explanation of the relative absence of initial positive ions under these conditions.

(*b*) *Air containing Alcohol Vapours.*—Some experiments were also carried out in air containing known concentrations of propyl alcohol vapour obtained, as in the work of Tyndall and Phillips,* from solutions of the alcohol in quinoline. In Table II the results for negative ions are compared with those obtained by the previous method of Tyndall and Phillips. The values of the mobility by the two slit method are calculated on the assumption that each method gives the same value in dry air.

Mol. percentage of propyl alcohol.	Estimated vapour pressure of alcohol. In mms. Hg.	Mobility. Tyndall and Phillips.	Mobility. Two slit method.
0·0	0·0	2·15	(2·15)
1·6	0·25	1·7	1·52
7·7	1·1	1·36	1·31
14·9	2·1	1·14	1·16
28·3	4·1	0·98	1·00
63·0	9·0	0·83	0·86

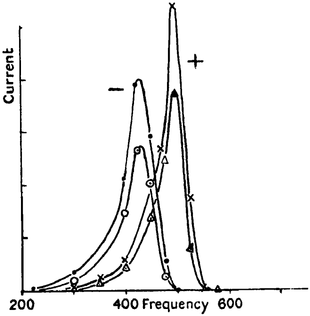


Fig. 6.—Air containing Propyl Alcohol Vapour.

- ● Short aged Negatives. × Short aged Positives.
- ⊙ Long aged Negatives. ▲ Long aged Positives.

---

* 'Roy. Soc. Proc.,' A, vol. 111, p. 577 (1926).

It was also found that within the limits of age investigated (*i.e.*, from 1/65 to 2/3 second, excluding the time of measurement), the mobility of both positive and negative ions in air containing organic vapours is independent of age. An illustration of this for air estimated to contain 9 mms. of propyl alcohol is given in fig. 6.

Summary.

Two methods of measuring mobility, both of which have a high resolving power, are described. In each method the apparatus may be made airtight, and the ions may be given any required age (within certain limits) before their mobility is measured.

The following experimental results have been obtained :—

(1) The rate of transformation of the positive ion in air is found to be higher by these methods than in earlier work, and there is some evidence that the transformation is a more complicated phenomenon than had been supposed.

(2) At long ages the positive ions in air have mobilities distributed over a small range with a mean value of about 1·25, and this mobility is independent of the humidity of the air.

(3) There is no evidence of initial positive ions in very dry air or in pure nitrogen. If any are formed, they nearly all transform in less than 1/100 second.

(4) The mobility of both the negative and positive ions in air containing alcohol vapour is independent of the age of the ions from 1/65 to 2/3 second.

The results given in this paper with regard to the transformation of positive ions are supplemented and discussed in Part V.

Our thanks are due to Mr. G. C. Grindley, who, indeed, suggested one of the methods employed, and to Mr. A. S. Hill for help given in the early stages of the investigation. Our thanks are also due to the Colston Research Society of the University of Bristol for a grant in aid.

The Mobility of Ions in Air. Part V.—The Transformation of Positive Ions at Short Ages.

By A. M. TYNDALL, G. C. GRINDLEY and P. A. SHEPPARD.

(Communicated by A. P. Chattock, F.R.S.—Received July 25, 1928.)

It was shown in Part IV of this paper that both the methods there described, which involved the use of closed vessels, gave variable results for the rate of transformation of positive ions, and values of the rate which were higher than those obtained by Erikson or by Tyndall and Grindley.*

It was felt that this might be due in part at least to products of chemical action arising from the α-rays, or to the impurities given off from the walls which would accumulate in a closed vessel unless some agent is present to remove them. To investigate this possibility further, the writers set up an air-blast method in which the effect of adding or taking away such impurities could be conveniently studied.

Apparatus.

The method is identical in principle with that by which Erikson first showed the existence at short ages of initial positive ions possessing the same mobility as negative ions. It differs from it, however, in detail, and except in dimensions and speed of air blast it resembles more closely that used by Grindley and Tyndall† for the investigation in one gas of ions produced in another. Fig 1 (A)

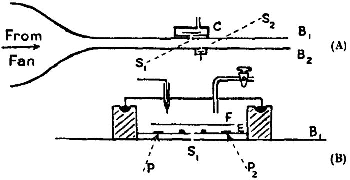


FIG. 1.

is a sketch of the arrangement. Air from a Sirocco blower converges into a narrow rectangular channel between two brass plates B_1 B_2, 0·7 cm. apart

---

* 'Roy. Soc. Proc.,' vol. 110, p. 342 (1926).

† 'Phil. Mag.,' vol. 48, p. 711 (1924).

The plates, separated by distance pieces of ebonite, are much closer than in Erikson's apparatus, thereby reducing the portion of the life of the ion spent in the measurement of the mobility itself to a very low value—about 1/1200 second.

Ions from polonium are generated in the box C: they are drawn by a subsidiary field through a narrow slit S_1 in B_1, and across the channel by the main field and down the channel by the air blast. For a given air blast the number reaching the slit S_2 in B_2 depends upon the strength of the main field, and the mobility of the ions. Those that enter S_2 are carried by a third field to the electrometer plate D. The box C is shown in more detail in fig. 1 (B). It is bounded by a circular wall of ebonite. Its base is the top plate B of the channel, and its lid is an inverted glass dish resting in a groove in the ebonite sealed with mercury. E is a brass plate 3 mms. from B_1, with a narrow slit in it vertically over S_1. Its function is to cut down the effect of eddies from the air stream at S_1 in the ionisation chamber.

Polonium at P_1 and P_2 ionises the narrow region 3 mms. deep between another plate F and E; small metal stops, shown as little black squares in the figure, further restrict the ions over the slit to a still narrower layer close to F. The age of all the ions entering S_1 is then the same, and appropriate fields between F, E and B fix its value. The field in the air channel may be varied from 600 to 1600 volts per centimetre, the critical region for initial ions being about 950.

Through the lid of the box a tube is inserted for the supply of any required gas to the ionisation chamber, and also a point from which a subsidiary discharge producing products of chemical action in air may be obtained.

Results.

Moist Air.—A typical result is shown in fig. 2, in which the electrometer current is plotted with voltage across the air stream. The voltage at which a peak in a curve occurs is *inversely* proportional to the mobility of the ions reaching the electrometer. Curve A is for negative and Curve B for positive ions. They are of the type found by Erikson. The curve for positives shows a number of initial ions of the same mobility as that of the negatives, but some final ions are also present, though not in sufficient numbers to give a sharp peak. The age in this case was about 0·005 second. At longer ages a pronounced peak was obtained for final ions, and the ratio of the mobilities of initials and finals was found to be 1·4, in agreement with Erikson's figure.

The rate of transformation was, however, invariably found to be higher than that given by Erikson. This seemed significant, because if the transformation is influenced by the presence of products of chemical action, this result follows from the use of a closed form of ionisation chamber in which these products can accumulate.

To study the matter further, the following series of experiments were carried out:—

1. *Dry Air.*—A slow stream of air dried by strong sulphuric acid was led through the chamber, and out at the slit S_1 into the main air blast. As might be expected, the number of ions emerging from S_1 was thereby somewhat increased, but it was found that the ratio of the number of initials to finals was changed so that the percentage of final ions was appreciably raised. This is in agreement with the previous results of Tyndall and Grindley, in which they showed that the rate of transformation increased as the air is made drier.

2. *Ozonised Air.*—A slow stream of room air was supplied to the chamber, but before entering it it was passed through an ozoniser. The results are shown in fig. 2. The curve for negatives in ozonised air is for clearness not shown.

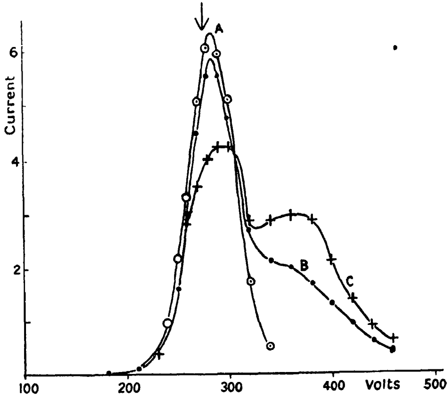


FIG. 2.—Moist Air.

● Positive in Room Air. ⊙ Negatives in Room Air.
+ Positive in Ozonised Air. ↓ Peak for Negatives in Ozonised Air.

It was practically identical with the air curve, but with its peak (indicated by an arrow) slightly displaced in a direction indicating if anything a slightly higher mobility in the presence of the ozone. The curve for positives shows definitely that ozonised air decreases the number of initial ions, and increases the number of final ions without producing any marked effect on their mobilities.

3. *A Subsidiary Point Discharge.*—A discharging point is similar to an ozoniser in its chemical action, and it therefore forms an alternative method of producing products in the ionisation chamber. In accordance with (2) it was found that the rate of transformation was increased by running a subsidiary point discharge in the vessel, the α-rays being still the source of the ions under investigation.

4. *Results with Ions from a Point of Discharge.*—In early work on the mobility of ions in air, Chattock* obtained the value 1·38 for positive ions from a discharging point. That is to say, this discharge gave final ions despite the fact that the maximum age of the ions could hardly have exceeded 1/500 second, and was probably of the order of 1/1000 second. To confirm this the polonium was removed and the upper slit replaced by a fine hole, with the point above it as the source of ions. The age of the ions was then of the order of 1/1000 second, or even shorter. The positive curve under these conditions had one peak only, coinciding neither with initials nor finals, but nearer the latter than the former. The relative absence of initials at this very short age is in agreement with Chattock's results, and is further evidence of the effects of products of chemical action on the rate of transformation.

5. *Oxides of Nitrogen.*—The products of chemical action in the above include both ozone and oxides of nitrogen. It seemed important to investigate their separate effects. For this reason the polonium and slit system was replaced, and a slow stream of air containing a small percentage of nitrous oxide led through the chamber. No detectable effect on the rate of transformation was found. The same result was obtained when a dish of copper and nitric acid was placed in the chamber, which then contained a certain amount of nitric oxide and nitrogen per oxide. But the reliability of the test was reduced by the rapid action of the fumes on the contents of the chamber and the polonium surface. It seems probable, however, that the effect produced by the ozoniser is not due to oxides of nitrogen.

6. *Ozonised Oxygen.*—The ozoniser was then replaced, but fed with oxygen instead of air. The effect of the ozone in this case was very striking, and is

---

* 'Phil. Mag.,' vol. 48, p. 401 (1899).

shown in fig. 3, in which A is the positive curve for a slow stream of oxygen and B that for the same stream ozonised. Clearly, the presence of ozone greatly increased the rate of transformation of initials into finals.

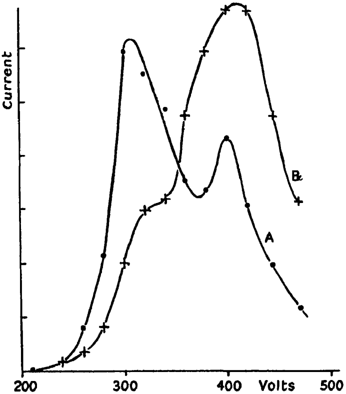


Fig. 3.—Ozonised Oxygen.

● Positives produced in Oxygen. + Positives produced in Ozonised Oxygen.

7. *Effect of Manganese Dioxide.*—In the presence of manganese dioxide the rate of decomposition of ozone is greatly accelerated. If trays containing it are placed in the ionisation chamber, the concentration of ozone produced by the α-rays in the enclosed air will be reduced, and one would therefore expect an increase in the relative number of initial ions due to a decrease in the rate of transformation. This expectation was experimentally confirmed. For instance, in one experiment the presence of manganese dioxide reduced the number of initials and raised the number of finals by about 25 per cent.

8. *Rapid Air Stream.*—One might expect a similar result from blowing a fairly rapid stream of air over the polonium so that the ozone produced by the α-rays is continuously removed. A small effect in the anticipated direction was obtained, but it was not as great as was expected.

9. *Nitrogen.*—A stream of fairly pure nitrogen from a cylinder was bubbled through a sulphuric acid solution to give it approximately room air humidity, and then passed through the ionisation chamber. Under these conditions very little ozone could have been present. The result is shown in fig. 4, in which A is the positive curve for an air stream and B for a nitrogen stream. Although, owing to a difference in the speeds of stream, the curves do not correspond in the areas they enclosed, it is quite clear that the percentage of finals is much greater in the nitrogen experiment. Ozone is, therefore, not the only factor which leads to the formation of final ions.

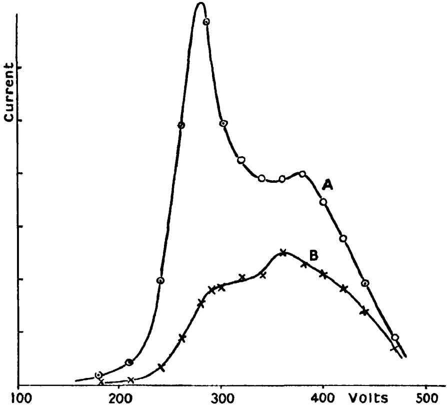


Fig. 4.—Nitrogen.

⊙ Air. + Nitrogen.

10. *Additional Results.*—Some additional experiments with the four gauze apparatus, described in Part IV of this paper, were then undertaken with a view to altering the concentration of ozone or other impurities present in the bell jar.

(*a*) A slow stream of air was circulated through the apparatus before and during the measurement. No definite effect was observed, though it may be argued that the stream was not fast enough in view of the relatively large volume of the apparatus.

(*b*) Electrodes were inserted so that a succession of sparks could be passed through the air before the measurements. In every case some increase in the number of final ions was observed.

(*c*) The effect of carbon dioxide in the air was investigated by comparing the results in room air, air bubbled through potash, and air with more carbon dioxide added to it. In air of humidity 8·2 mms. no effect due to carbon dioxide was observed In air of humidity 0·6 mm. the addition of carbon dioxide produced apparently a slight rise in mobility of the positives.

11. *Alcohol Vapours.*—Some further results were obtained with the air-blast method by forming the ions in air containing various alcohol vapours. Their mobility was measured as before in room air. The results for methyl alcohol and propyl alcohol are shown in fig. 5, the vapour in each case being

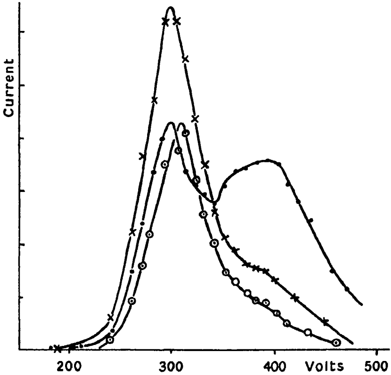


Fig. 5.—Positive Ions.

● Air. ⊙ Air with Methyl Alcohol. × Air with Propyl Alcohol.

approximately saturated from an open dish of the liquid in the chamber. In each case it is seen that the effect of the alcohol is the reverse of that of ozone, in that the final positive ions are very greatly reduced in number. The negative curves are for clearness not included, but they are identical whether the alcohol is present or not.

Discussion of Results.

It is important to note that there is one essential difference between this method and those used in Part IV of this paper. In the present air-blast method the ions in a number of cases are produced in one gas (ozonised air, nitrogen, alcohol vapour and air, and so forth), and measured in another gas (room air). From some earlier experiments on similar lines, Grindley and Tyndall* concluded that in this case the mobility is that of the ions of the gas in which the measurement is made, and their results are explained by assuming that the cluster of molecules of the original gas round the charge is shed and replaced by those of the air stream. The conditions, however, in the present experiment differ from these in that the time the ions spend in the air stream is of the order of 1/1200 second instead of about ½ second, and unless the time taken by the interchange is small compared with this, the effect of the initial gas molecules on the mobility should be observable.

If we consider first the results in air containing alcohol vapour, it is clear that there is strong evidence that the negative ions shed their attached alcohol molecules immediately they enter the air stream, as, indeed, for statistical reasons one would expect them to do. If it were not so, one would expect in the case of propyl alcohol of this concentration a reduction in mobility of more than 50 per cent.

The positive results suggest that in this case also the alcohol molecules are shed because in propyl alcohol of this concentration the mobility of the positive ions is at least 25 per cent. lower than that of the final air ion. It looks, therefore, as if owing to the fairly high concentrations of these vapours in the chamber, and the marked dipoles of their molecules, they had successfully competed for places at the surface of the ion at the expense of whatever substance is responsible for the transformation into finals. On emerging into the stream the alcohol molecules are knocked off by air molecules, leaving the initial ions, which cross the air stream in 1/1200 second before they have had time to take up the form which gives them their final character.

Other experiments in the series may be regarded as showing that some of the chemical products of ionisation in air (probably ozone) increase the rate of transformation. Indeed, it seems probable that the discrepancies between the previous results of Chattock, of Erikson and of Tyndall and Grindley, together with those by the four gauze and two slit methods, may be explainable on these lines. Thus it may be noted that in these methods the rate of transformation seems to follow the probable concentration of products present. The maximum rate of transformation is found in Chattock's work, where the quantity of products are also high The closed vessel methods of this paper come next, while the air-blast method of Erikson and the flash method of Tyndall and Grindley, in which the products are at a minimum, give the slowest rate of transformation The fact, taken alone, might suggest that the final ion in air is ozone; but the results in pure nitrogen, given in Part IV, show conclusively that this is not the case. On the whole, the results suggest that the initial ion depends for its formation on the presence of water vapour, and is transformed to a final stable ion on collision with an ozone molecule. The relatively slow rate of transformation may possibly be due to the dependance of the change on some type of three body collision. Whether the transformation would occur without contact with ozone is a moot point.

---

* 'Phil. Mag.,' vol 48, p. 711 (1924).

In this connection there is one fact which may be significant. Ozone is relatively stable in air at room temperature, an appreciable quantity remaining in a closed vessel after many hours. Perman and Greaves* showed, however, that its rate of decomposition is accelerated by water vapour. Methyl alcohol, judging from its chemical properties, is even more active. The increase in initial ions that these substances seem to promote might, therefore, be not directly due to them, but due to the reduction they bring about in the ozone concentration.

Summary.

1. The effect on the rate of transformation of positive ions of adding certain gases and vapours to the air in which they are formed has been studied by an air-blast method.

2. It is shown that a small quantity of ozone, generated either by an ozoniser or by a point discharge, produces a marked increase in the rate of transformation; indeed, there is evidence that sufficient ozone may be generated by α-rays from polonium in a closed vessel to produce an appreciable effect. It is argued that this effect may explain certain discrepancies between the results of mobility measurement made by various observers using different methods.

3. The rate of transformation is greatly retarded if, before entering the air

---

* 'Roy. Soc. Proc.,' vol. 80, p. 353 (1908).

blast, the ions are formed in an atmosphere containing the vapour of certain alcohols of the aliphatic series.

4. The significance of these results is discussed, but at this stage speculation regarding the nature of the initial positive ion appears to be premature.

Our thanks are due to the Colston Research Society of the University of Bristol for a grant in aid of this work

---

A Mechanical Method for solving Problems of Flow in Compressible Fluids.

By G. I. Taylor, F.R.S., and C. F Sharman, Ph.D.

(Received August 28, 1928.)

[Plate 2.]

Introduction.

At the present time the chief study of aerodynamical laboratories is concerned with the steady flow of air past solid bodies at speeds which are so low that the effect of compressibility is inappreciable. In recent years, however, the rapid increase in the speed of aircraft has very much increased the importance of the study of the effect of compressibility on air flow. The highest speeds of aircraft at the present time are of the order of 280 miles per hour or 400 feet per second, *i.e.*, 0·4 of the speed of sound. The tips of the propellers of these high speed machines may move at speeds as high as 1·3 times the speed of sound. At low speeds when air behaves like an incompressible fluid, the classical theory of hydrodynamics which is concerned with irrotational motion predicts that a body moving steadily will experience no resistance or lift. In the simpler form of the theory of Prandtl the motion is still irrotational, but the existence of a circulation round the body and of vortex sheets trailing down-wind permit the possibility of resistance and lift.

At speeds higher than that of sound it is known that a nearly discontinuous wave is formed in front of the body. This wave has been photographed and has been studied theoretically by many writers. It involves a dissipation of energy so that even on the classical theory which involves no viscosity, bodies moving at speeds higher than that of sound should have a resistance.

At speeds less than, but comparable with, that of sound it seems clear that if the effect of viscosity is neglected and if it is assumed that no discontinuous wave is formed the classical theory of hydrodynamics will predict that the resistance will be zero. This has in fact been proved by Cissotti and by Rayleigh.

It has been assumed by most writers that the problem presented by the steady motion of a body through a compressible fluid has a solution in which the flow is irrotational provided that the speed of the body is less than that of sound. Making this assumption Rayleigh* and Glauert† have attempted to find how the flow round a body placed in a stream of air is modified by compressibility.

In both cases the difficulties of the analysis forced them to restrict their problems to such an extent that they were unable to find out anything about the conditions under which irrotational motion becomes impossible. In Glauert's analysis it is assumed that the velocity nowhere differs appreciably from the undisturbed velocity. This limits his discussion either to the flow at a great distance from the body, or to the flow due to a narrow body, like a thin aerofoil, the surface of which makes only a small angle with the undisturbed flow. And even then if the thin aerofoil is set at such an angle that it produces a velocity appreciably different from that of the main stream Glauert's analysis ceases to be applicable.

Rayleigh's analysis is restricted to the case of flow past a circular cylinder. He obtained a first approximation giving the modification which incompressible flow experiences owing to compressibility. His analysis, so far as he worked it out himself, is subject to the limitation that the speed of the stream must be small compared with that of sound. On the other hand, it is theoretically capable of being applied to find higher approximations, but the mathematical difficulties increase rapidly with each successive application and even the second approximation is almost unmanageable. As a result of his consideration of the form which the solution would ultimately assume if he could carry his method of successive approximation far enough, Rayleigh came to the following general conclusion: "As regards the general question it would appear that so long as the series are convergent there can be no resistance and no wave as the result of compressibility. But when the velocity V‡ of the stream exceeds that of sound the system of velocities in front of the obstacle expressed by our equations cannot be maintained, as they would be at once swept away downstream. It may be presumed that the passage from one state of affairs to the other synchronises with a failure in convergency. For a discussion of what happens when the velocity of sound is exceeded reference may be made to a former paper."*

---

* 'Phil. Mag.,' vol. 32, p. 1 (1916).

† 'Roy. Soc. Proc.,' A, vol. 118, p. 113 (1928).

‡ In Rayleigh's analysis V is the velocity of the stream at a great distance from the body.

It will be seen that Rayleigh considered that the solution involving irrotational motion would hold till the speed of the main body of the stream reached that of sound, and that at that speed a failure of convergency in his successive approximations would indicate the formation of a discontinuous wave of the type which is known to exist in front of bodies moving at speeds higher than that of sound.

Owing to the difficulty in carrying out the successive approximations Rayleigh appears to have been unable to verify this presumption, and even had he been able to do so the result would have been applicable only to the case of a circular cylinder.

In the following pages a method is described by which it is possible, without very great labour, to find the flow of a compressible fluid past an obstacle of any shape provided that irrotational motion is possible. The first application of this method has been to Rayleigh's case of the circular cylinder. In this case the method produces, mechanically, successive approximations which are practically identical with those which would result from successive applications of Rayleigh's mathematical method. One of the most important results of the present work is the discovery that the failure of convergence occurs not when the speed of the main body of the stream reaches that of sound but at some lower speed. In the case of a circular cylinder convergence fails *when the maximum velocity in the field reaches the speed of sound in the air at that point*. This first occurs when the speed of the stream is between 0·4 and 0·5 of that sound. *At a speed of* 0·5 *of that of sound it appears that no continuous irrotational motion is possible* past a circular cylinder.

The method is applicable to cylindrical bodies of any cross section and even to bodies like aerofoils which carry a circulation with them. One of the principal results to be anticipated from further work is the determination of the maximum speed at which irrotational motion is possible with bodies of various shapes.

---

* 'Roy. Soc. Proc.,' A, vol. 84, p. 247 (1910).

Bryan's Method.

An attempt to solve problems in compressible flow when the limitations of Rayleigh's and Glauert's analysis are not imposed was made some years ago by Bryan.* Starting with the flow of an incompressible fluid (supposed known) past a cylindrical obstacle of any given cross-section Bryan divided the field into squares by means of a network of stream lines and lines of equal velocity potential. He then distorted these squares into rectangles in accordance with formulæ depending on the law of adiabatic compression and attempted to construct a new pattern of stream lines and equipotential lines to represent the flow of a compressible fluid. To complete the solution of any problem it was necessary to make these rectangles fit into a pattern in which stream lines and equipotential lines cut at right angles, and at the same time to make them fit the boundary of the solid body, and become straight at great distances from the cylinder. The difficulty involved in doing this seems to have been so great that the method proved a failure. At any rate, no concrete problem appears to have been solved by it.

Theory of Electrical Analogy.

It appears that directly the limitations imposed by Glauert's or Rayleigh's methods are dropped the amount of labour involved in solving, by purely analytical methods, any case of flow of a compressible fluid past an obstacle is so great that there is little hope that any useful results will be obtained in this way. For this reason it seemed worth while to develop an analogy which was discovered by one of us between the flow of a fluid of variable density in two dimensions and the flow of electric current in a conducting sheet of variable thickness. Let us compare the equations which represent the irrotational flow of a fluid in which the density ρ is a function of two co-ordinates x and y, with the flow of current in a sheet of small variable thickness t (a function of x and y) and constant specific resistance σ. The equations of hydrodynamic flow are

$$\left.\begin{aligned} \frac{\partial\phi}{\partial x} &= -u, & \frac{\partial\phi}{\partial y} &= -v \\ \frac{\partial\psi}{\partial x} &= \rho v, & \frac{\partial\psi}{\partial y} &= -\rho u \end{aligned}\right\}, \tag{1}$$

---

* 'Reports and Memoranda Adv. Committee for Aeronautics,' 1918-19, p. 130, and 1919-20, p. 79.

where ϕ is the velocity potential, ψ is a stream function and u, v are the components of velocity.

The equations of electric flow are

$$\left.\begin{aligned} \frac{\partial V}{\partial x} &= -f\sigma, & \frac{\partial V}{\partial y} &= -g\sigma \\ \frac{\partial W}{\partial x} &= tg, & \frac{\partial W}{\partial y} &= -tf \end{aligned}\right\}, \tag{2}$$

where V is the electric potential and W is an electric current function. f and g are components of current density.

Comparing equations (1) and (2) it will be seen that they are identical in two cases, either

$$\text{(A) when } V = \phi, \ W = \psi, \ u = f\sigma, \ \rho v = tg, \ v = g\sigma, \ \rho u = tf$$

and these are consistent if $t = \rho\sigma$ (3)

or (B) when $W = \phi, \ V = -\psi, \ -u = tg, \ \rho v = f\sigma, v = tf, \ \rho u = -g\sigma$

and these are consistent if $t = \sigma/\rho$. (4)

It appears therefore that if we wish to make use of a mathematical analogy to find the distribution of flow in a fluid of variable density we can do so in one of two ways.

(A) We can construct a conducting sheet the thickness of which is variable and is everywhere proportional to ρ. In this case the lines of equal electric potential are lines of equal velocity potential and lines of flow of the electric current are the stream lines of the fluid. The intensity of electric force is proportional to the velocity of the flow and the direction of greatest electric intensity is parallel to the direction of flow.

(B) If the thickness of the conducting sheet is made inversely proportional to ρ the equipotential lines of the current distribution are stream lines and the lines of electric flow are lines of equal velocity potential in the field of fluid flow. In this case the product of the electric intensity and the thickness of the sheet is proportional to the velocity of the stream, and the direction of greatest electric intensity is perpendicular to the direction of the stream in the field of fluid flow.

Boundary Conditions.

To represent a uniform flow either in (A) or (B) it is necessary to produce a uniform field. For this purpose it is convenient to use a rectangular sheet of small conductivity. The current is fed into this by means of large copper strips in contact with opposite edges of the sheet. To represent the flow past a solid cylindrical body with any given cross-section it is necessary in case (A) to cut a hole in the sheet of the same shape as the given cross section. There is then no flow of current across the boundary, and this corresponds with the boundary condition in the field of fluid flow that there shall be no flow of fluid across the surface of the solid.

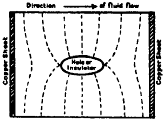


Fig. 1.—Analogy (A).

The appropriate sheet for investigating the flow past an elliptic cylinder with its long axis parallel to the direction of the stream is shown in fig. 1. The equipotential lines are shown as broken lines and the stream lines are perpendicular to these. If it is desired to use analogy (B) the conducting sheets must be placed along the edges which are parallel to the stream and the section of the solid cylinder must be represented by a perfectly conducting lamina or disc. The appropriate electric arrangement for representing the same flow as before is shown in fig. 2. The equipotential lines which are also fluid stream lines are shown, and the perfectly conducting lamina is represented by the ellipse E.

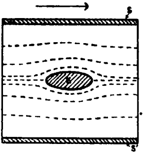


Fig. 2.—Analogy (B).
Flow without Circulation.

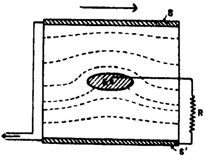


Fig 3.—Analogy (B).
Flow with Circulation.

Representation of Circulation.

If the disc E is connected electrically with the sheets S and S′ (fig. 2) only through the conducting sheet, then the total amount of current which flows into E is equal to the amount which flows out. This condition is represented mathematically by the equation

$$\int\left(-ft\frac{dy}{ds}+gt\frac{dx}{ds}\right)ds=0, \tag{5}$$

the integral being taken round the edge of E. Making use of the analogy (B) this becomes

$$\int\left(-v\frac{dy}{ds}-u\frac{dx}{ds}\right)ds=0. \tag{6}$$

In the field of fluid flow (6) is the condition that there shall be no circulation round the body.

To represent irrotational motion with circulation it is necessary to lead some current into the lamina E by an external wire so that the integral in (5) is not zero.

In fig. 3 is shown the arrangement for representing a circulation. The lamina E is connected with one of the edge plates, say S′, by an external resistance R, and by adjusting R any given amount of circulation can be represented in the electric field. The stream lines are sketched in fig. 3.

It is not possible to represent a fluid flow with circulation by means of analogy (A) because in that case the electric potential would have to be many-valued, and it is not possible to realise this condition as a steady state.

Introduction of the Gas Equation.

The analogies (A) or (B) could be used directly to determine the stream lines in any two-dimensional field of flow for which the density is known as a function of position.

In the case of a compressible fluid, like air, the density is known as a function of pressure through the gas equation, and the pressure is known as a function of density and velocity through Bernouilli's equation, so that the density is known as a function of velocity. It is clear that if by some means we could find the density at every point we could use either of the analogies to test whether our solution was correct, for we could construct the appropriate conducting sheet, and we could measure the electric intensity at every point. From the electric intensity we could find the corresponding velocity by making use of the analogy, and from the gas equation and Bernouilli's equation we could find the density corresponding with this velocity. If this density was the same as the original assumed density at all points, then the original assumed density was correct.

If our original assumed density is not correct, we can still carry out the process outlined above and obtain a new density distribution which is not the same as the original one. Starting with a new density, we can again apply the same series of operations and find a third density distribution. This iterative process can be repeated indefinitely, and if it is found that the difference between one density distribution and the next becomes less and less as the number of applications of the process increases, the method gives a converging series of approximations to the true solution of the problem.

Rayleigh's Method.

The method proposed by Lord Rayleigh for solving problems of two-dimensional flow in a compressible fluid may now be compared with the new method proposed above.

The equation for the velocity potential of a fluid of variable density may be expressed in the form

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = -\frac{1}{\rho}\left(\frac{\partial\rho}{\partial x}\frac{\partial\phi}{\partial x} + \frac{\partial\rho}{\partial y}\frac{\partial\phi}{\partial y}\right). \tag{7}$$

Using Bernouilli's equation, $\int\frac{dp}{\rho} + \frac{1}{2}q^2 = \text{constant}$, where $q^2 = u^2 + v^2$.

(7) becomes

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = \frac{1}{2\,\dfrac{dp}{d\rho}}\left(\frac{\partial\phi}{\partial x}\frac{\partial q^2}{\partial x} + \frac{\partial\phi}{\partial y}\frac{\partial q^2}{\partial y}\right) \tag{8}$$

and

$$\frac{dp}{d\rho} = a^2 - \frac{1}{2}(\gamma - 1)\,q^2,$$

where a is the velocity of sound in the fluid at some distance from the body and γ is the ratio of the specific heats.

Lord Rayleigh began by considering the case when the effect of compressibility is small, so that the flow is not very different from that of an incompressible fluid. He assumed that the solution was known for incompressible flow when $\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$.

He then took the values of $\frac{\partial\phi}{\partial x}, \frac{\partial q^2}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial q^2}{\partial y}$ found for incompressible flow and inserted them in the right-hand side of (8). In the case where the speed of flow is small compared with that of sound $dp/d\rho$ is constant to the order of approximation required, and equal to the square of the velocity of sound. The method consists in finding the solution to the equation

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = \frac{1}{2a^2}\left(\frac{\partial q^2}{\partial x}\frac{\partial\phi}{\partial x} + \frac{\partial q^2}{\partial y}\frac{\partial\phi}{\partial y}\right), \tag{9}$$

where the right-hand side of this equation is given the value which it has for an incompressible fluid. The result gives a new solution which contains terms having v^2/a^2 as a factor, as well as the terms which occur in incompressible flow. *This first approximation to the effect of compressibility was worked out by Lord Rayleigh for the case of flow past a circular cylinder. He states that the value of ϕ so obtained can be inserted in equation (8) and the solution used to obtain a second approximation.* It must be remembered, however, *that to obtain the second and higher approximations q^2 cannot be neglected* compared with a^2, and $dp/d\rho$ which is the square of the local velocity of sound differs appreciably from a^2. This makes the labour involved in obtaining higher approximations so great that the method becomes impracticable. Rayleigh considered, however, that if it could be carried out the method would give a converging series provided that V, the velocity of the uniform stream, were less than a, the speed of sound in the undisturbed stream.

Comparison between Present Proposals and Rayleigh's Method.

The electric analogy method now proposed amounts really to finding successive approximations to (7) using a value of ρ derived from the previous approximation in the same way that Rayleigh derived the whole of the right-hand side of (8) from the previous approximation. It will be seen that the two methods are nearly identical in theory though very different in the manner in which they are applied. The chief difference is that the terms $\partial\phi/\partial x$, $\partial\phi/\partial y$ on the right-hand side of (7) are taken from the previous approximation in Rayleigh's case, whereas in the proposed method they are connected with the new value of ϕ found by means of the analogy.

The advantages of the proposed method are (*a*) that it is no more difficult to derive the $(n+1)^{\text{th}}$ approximation from the n^{th} than it is to derive the second from the first.

(*b*) That it is no more difficult to apply the method to a body of any shape (an aerofoil for instance) than it is to apply it to the case of a circular cylinder

Auxiliary Curves for applying Gas Equations.

The analogy is used to find values of q appropriate to the distribution of density assumed as the result of a previous application of the method. This can most conveniently be expressed by the symbol θ where $q = \theta V$, V being the velocity of the uniform stream in which the body is immersed. If a is the velocity of sound in the fluid at the temperature and pressure which it possesses in the undisturbed stream, and if $n = V/a$, then combining Bernouilli's equation with the gas equation $p\rho^{-\gamma} =$ constant, it will be found that

$$\left(\frac{\rho}{\rho_0}\right) = \left\{1 - \tfrac{1}{2}(\gamma - 1)\,n^2(\theta^2 - 1)\right\}^{\frac{1}{\gamma-1}}, \tag{10}$$

where ρ_0 is the density in the undisturbed stream. To make use of (10) in any particular case it is convenient to have curves giving the value of ρ/ρ_0 in terms of θ for several different values of n. Three of these, for $n = 0 \cdot 4$, $0 \cdot 5$ and $0 \cdot 6$, and $\gamma = 1 \cdot 407$* are shown in fig. 6.

Apparatus.

The most convenient substance for making a conducting sheet of variable thickness is an electrolyte such as copper sulphate solution contained in a tank of small but variable depth. It was found possible to construct a tank in which a bottom made of paraffin wax could be carved without difficulty into the shape required. Details of the technique developed for this purpose are described later.

A tank of uniform depth has been used at the National Physical Laboratory for determining the flow of incompressible fluid. In this tank the electrolyte was dilute copper sulphate solution, and the body round which the flow was to be explored was represented by a copper cylinder cut to the correct section. An alternating current passed between copper sheets covering opposite sides of the tank, and pairs of points having the same potential were found by dipping into the electrolyte two exploring electrodes consisting of copper wires connected through a telephone. When the two wires were at points on the same equipotential line no buzzing should be heard in the telephone. When using telephones to detect an alternating current it is not always possible to eliminate altogether the effects of induction, so that there may be no point at which the telephone is quite silent. For this reason we adopted a suggestion of Mr. E. F. Relf, and used an alternator to supply alternating current across our tank. The current from the exploring points was rectified by a commutator placed on the same shaft as the armature of the alternator. It was then detected or measured by a sensitive galvanometer.

In the apparatus used at the National Physical Laboratory stream lines were drawn by fixing one of the exploring electrodes at some point near the edge of the tank and moving the other along the equipotential line passing through that point. The velocity was determined by graphical differentiation of the resulting pattern of stream lines. In the present work we were not so much interested in the form of the stream lines as in the magnitude of the velocity at each point. For this reason we used an exploring system consisting of two electrodes placed at a fixed small distance apart, and mounted in a carrier capable of rotation about a vertical axis passing mid-way between them.

---

* The ratio of the specific heats for air is $\gamma = 1 \cdot 407$.

With this system direct measurements of the direction and magnitude of the electric field could be made. In each case the ratio of the electric field at the point concerned to the uniform electric field at great distances from the body was measured. In the case of analogy (A) this is equal to θ and in the case of analogy (B) it is equal to $\rho\theta$.

If analogy (A) is being used the value of θ obtained in this way is inserted in equation (10), and the corresponding value of ρ/ρ_0 found for each point of the field. A new bottom is then cut for the tank, so that the depth of the fluid at each point is ρ/ρ_0 times the depth at points corresponding with the uniform stream, and another set of measurements of electric field are made. These form the basis for yet another approximation. As soon as one approximation is sensibly identical with the previous one the problem may be regarded as solved.

Detailed Description of Apparatus.

A liquid electrolyte having been decided upon for the material of the current sheet, the following conditions had to be satisfied :—

(*a*) The electrolyte must be contained in a large shallow water-tight tank composed throughout of electrical insulators, and the bottom must be of a substance which is easy to model. At the same time appreciable distortion of the bottom by bending must be eliminated by rigid construction and careful suspension.

(*b*) The tank requires steady supports and must be capable of accurate levelling. The importance of levelling lies in the fact that the current sheet must be thin in order to comply with the condition of two-dimensional flow, and hence a small error in levelling would produce a large percentage variation of thickness across so large a surface.

(*c*) For the location of the exploring points and for the modelling of the bottom of the tank a three-dimensional Cartesian co-ordinate system is necessary, and here again great rigidity is essential, especially in the vertical direction.

In order to satisfy these requirements the tank and the co-ordinate system were both built up from a slab of slate (AA in fig. 4), which was 4 feet square and 1½ inches thick. This was cut, and its upper surfaces rubbed plane to within ½ mm. all over,* and it was considered by the makers to be rigid enough to remain so under its own weight on a three-point support.

---

* By Messrs. E. J. Riley, Accrington.

A carefully joined framework of 1-inch teak, 5 inches high and 36 inches by 38 inches inside, formed the sides of the tank, two of which are shown in section at BB in the figure. This was made waterproof by ironing into it paraffin wax and was clamped tightly down to the slate by four stout clamps (not shown). The joint between the sides and the slate was sealed inside with a fairly hard " vacuum " wax, but it was found necessary after a time to lute the outside with " plasticine."

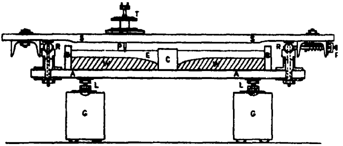


Fig. 4.

The co-ordinate system consisted of two pairs of parallel steel rails, of which one pair SS formed a rigid carriage sliding on the other pair RR and perpendicular to it. A second carriage T, made from a brass casting slid on the upper rails and carried the depth measuring arrangement or the exploring points. The rails were four circular steel rods, $1\frac{1}{4}$ inches in diameter and $4\frac{1}{2}$ feet long, which were tested for straightness by rolling on the slate surface. One pair, shown in section as RR in figure, was fixed to the slate by means of six steel pillars, which had been carefully made the same height. The bottom of each pillar had a turned flange and the top a rectangular groove, into which the rails were pulled by a pin and bolt arrangement. The coplanarity of these two rails should be as good as that of any two lines drawn in the slate surface ; their parallelism was adjusted by means of a distance piece.

The great length of the first carriage, SS, caused the idea of a geometrical slide to be abandoned, for a slight deviation from straightness of the fixed rails would involve a large error in the placing of the exploring points. The carriage was supported by four brass buttons of equal thickness screwed to the upper and sliding on the lower rails. Horizontal motion in a direction perpendicular to the fixed rails was limited by a fixed button at one end and a spring-loaded one with a clamping screw at the other. Pointers were attached to both ends of this carriage reading in scales fixed to the lower rails.

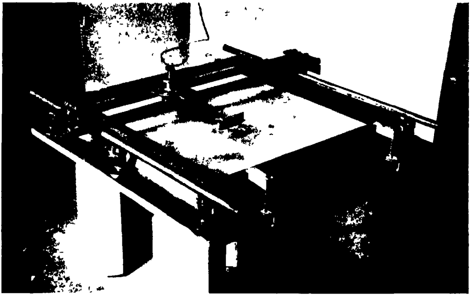
The smaller carriage, T, which was entirely of brass, had a geometrical slide and its position was read on a single scale. It carried a graduated turn-table whose axis was accurately perpendicular to the surface of the slate. The brass tube containing the exploring electrodes passed coaxially through the turn-table and was fixed by a split collar. For measurements of depth and for levelling a graduated steel rod furnished with a sharp point was substituted for the electrode tube.

The slate slab was supported finally by five ⅝-inch steel levelling screws on concrete pillars—four at the corners of a rectangle and one at the centre. In order to level the tank and the whole co-ordinate system solution was poured into the tank, and the centre screw and one other were screwed down out of contact with the slate. Then, when the fluid surface was steady, the steel point was lowered so as nearly to touch it, and the remaining three screws adjusted until the distance between the point and its image was constant over the whole tank. The other screws were then turned up to take a small part of the load. From this procedure it was estimated that the point did not deviate more than 1/10 mm. from a horizontal plane anywhere in the tank.

The normal depth of the electrolyte, *i.e.*, the depth at "infinity," was 2 cm., which could be accurately checked by means of the graduated rod.

The wax bottom, of which the normal depth was 5 cm., was built up by pouring in successive layers of molten paraffin, each about ½ cm. deep, being allowed to solidify before applying the next. Air bubbles were removed by playing a Bunsen burner over the surface of every layer. The obstacle in these experiments was a circular cylinder of hard wood, which had been soaked in paraffin wax, 8 cm. long and 10 cm. in diameter. It was not convenient to have this in the tank while the bottom was being modelled, and so a hole was left in the wax and the cylinder was inserted when the bottom was finished. For this purpose a hollow brass cylinder of the same diameter was used as a mould for the hole, and it was removed by pouring in boiling water when sufficient wax for modelling had been run on to the bottom.

The wax was first cut roughly to shape with a knife. Then the depth for a given point in the field was calculated from the electric intensity measured at that point in the previous experiment. The steel point was next fixed at this depth in the slide, T (fig. 4), and the carriage was moved sideways along its slide till the steel point reached its correct position. This process produced a scratch in the wax which had its greatest depth at the point in the field corresponding with the calculated depth. When enough scratches of this kind had been made the wax was scraped with a scraper till the scratches



just disappeared. The surface obtained was very good, any air bubbles which appeared were easily removed by heating with a Bunsen burner, adding a little molten wax and rescraping. The art of choosing the best positions for the measurements of electric intensity so that there should be no ambiguity in cutting the wax surface for the next approximation was quickly acquired.

The tank with its slides and rotating holder for the pair of exploring electrodes is shown in the photograph, Plate 2. The body seen projecting through the wax bottom in this photograph is a model of an aerofoil used in some later experiments.

The electrical arrangement is shown diagrammatically in fig 5. Plates of polished sheet copper, PP, were clamped to two opposite sides of the tank, and an alternating potential difference was applied to them from a small rotary converter. This machine was an old four-pole fan motor which ran normally at 350 r.p.m. Slip rings were fitted over the commutator from which an alternating current of 0·3 amp. could be taken at 50 volts. With a 0·5 per cent. copper sulphate solution in the tank the total current was about 0·2 amp. In the figure, D represents the motor commutator, A the slip rings connected directly to the plates PP, and S the synchronous commutator which rectifies the small alternating P.D derived from the exploring electrodes E in the tank. O represents the obstacle, in these experiments a circular cylinder.

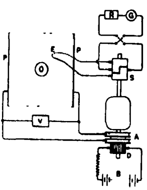


FIG. 5.

At first sight it might appear that a potentiometer would be the most suitable arrangement for measuring the small potential differences between the exploring points because there would then be no current flowing through them. It was found more convenient, however, to use a voltmeter, consisting of a sensitive gavanometer, G, in series with a high resistance R. In order that the disturbance in the electric flow due to the current taken by the voltmeter should be as small as possible, the resistance, R, should be large compared with the resistance of the electrolyte between the exploring electrodes. This condition was satisfied by putting 100,000 ohms in series with the galvanometer, and by using flattened wires, 1 mm. wide and 2 mm. apart, instead of wires of circular section as the electrodes, thus reducing the electrolyte resistance to about 100 ohms. The galvanometer circuit was completely insulated. It should be noticed that the small current through the electrodes was alternating, thus avoiding trouble due to polarisation there. Also the four contacts of the commutator S, which were made from the same sample of brass as the rotating sectors and were lubricated with pure paraffin, gave no disturbing potential differences.

The procedure adopted for the measurement of the velocity at any point in the field was as follows: The height of the surface of the electrolyte was first adjusted carefully by means of the graduated rod; then the exploring electrodes were fitted in the carriage and placed at a standard point in one corner of the tank—35 cm. from the centre along each axis. The turn-table was rotated until the galvanometer read zero and the angular reading was taken. It was next turned through a right angle to the position of maximum galvanometer reading, and the current supply to the generator was adjusted so as to give a galvanometer reading of 100 divisions. The constant running of the machine (the current for which was taken from a large storage battery, B, fig. 5) was checked by an A.C voltmeter (V, fig. 5) connected across the tank. The carriage was now moved to the point under investigation, turned to give zero reading, turned through 90° and the galvanometer read. In the analogy (A), the ratio of this reading to the standard reading of 100 divisions is equal to the ratio of the velocity at the point to the velocity at the standard point; and the standard point was sufficiently far removed from the cylinder that it could be regarded as at infinity.

Application to the Flow past a Circular Cylinder.

The first application of the method was to a circular cylinder placed in a stream of air moving with a speed of 0·4 times that of sound, so that $n = 0 \cdot 4$. Analogy (A) was used, the cylinder being represented by a turned wood block 10 cm. diameter. The bottom was first cut flat and horizontal so that the depth of electrolyte was uniform. The electric field in this case is known, and the first set of measurements corresponding with flow of an incompressible fluid were compared with the calculated values in order to test the apparatus. It was found to work satisfactorily and to give results correct to 1 per cent., due allowance being made in the calculations for the small effect of parallel walls in the position of the copper plates, PP (fig 5).

A new bottom was next constructed according to the principles set forth in the preceding section. The depth of the electrolyte at points along the line through the centre of the cylinder at right angles to the stream is shown in fig. 8.

The positions of points where some of the measurements were taken are shown in fig. 7 Owing to symmetry it is only necessary to take readings in one quadrant. The points are marked 1 to 16 in fig. 7. Their co-ordinates and values of θ taken with three successive bottoms to the tank are given in Table 1. The column headed θ_1 contains calculated values of θ for a flat bottom (some measurements were made with a flat bottom, and these agreed with the figures here given). These values of θ were transformed into values

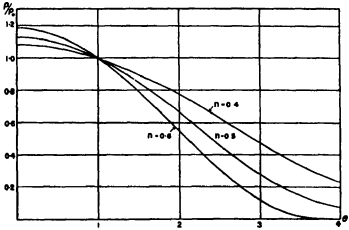


FIG 6.

of ρ/ρ_0 by means of the curve shown in fig. 6, and a new bottom was cut so that the depth of the electrolyte should be proportional to these.

Measurements of the electric intensity at the 16 points of Table 1, and also

Table I.—Values of θ in successive Approximations when $n = 0·4$.

Number of point in in fig 7.	x	y	θ_1	θ_2	θ_3
	cm	cm			
1	0	5·2	1·95	2·13	2·13
2	0	6·0	1·69	1·89	1·89
3	0	7·5	1·45	1·60	1·60
4	0	10·0	1·25	1·36	1·35
5	2·5	5·0	1·61	1·70	1·72
6	2·5	10·0	1·21	1·32	1·31
7	5·0	5·0	1·11	1·16	1·14
8	5·0	10·0	1·13	1·19	1·19
9	10·0	10·0	1·00	1·04	1·02
10	7·5	5·0	0·99	0·98	0·98
11	5·0	2·5	0·83	0·84	0·84
12	10·0	5·0	0·90	0·90	0·90
13	10·0	0	0·73	0·76	0·75
14	7·5	0	0·56	0·58	0·58
15	6·0	0	0·34	0·31	0·32
16	5·0	0	0*	0*	0*

* Estimated from measurements made as close to the surface of the cylinder as possible.

at several others, were next taken and the ratios of these to the readings at some standard position in the uniform field outside the range of influence of the cylinder were found. These ratios are given in the column headed θ_2 in Table I.

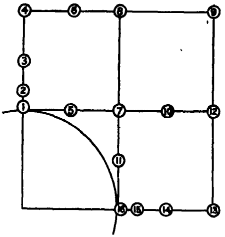


FIG. 7

The whole process was then repeated, using the values of θ_2 to find a new depth for the tank, and in this way a new set of values for θ were found. These are given in the column headed θ_3.

It will be seen that the convergence is very rapid, in fact θ_3 is so near θ_2 that it would have been impossible to detect any further change if the process had been repeated a fourth time. The figures given as θ_3 represent therefore the solution of the problem of compressible flow in the case when a circular cylinder moves at a speed equal to 0·4 times the velocity of sound. The effect of compressibility can be seen by comparing the fourth and last columns of Table I. It is to increase the speed nearly everywhere.

The depths of the electrolyte over the three successive bottoms are shown in fig. 8, where the curves represent depths along the section at right angles to the general direction of the stream.

The shape of the bottom can be visualised by means of the contour diagram shown in fig. 9.

These curves were found in the tank as curves of constant potential gradient, and the number attached to each is the ratio of the velocity at any point on it to the velocity at infinity. They are also curves of constant depth, but regarded as contours of depth they are not equally spaced. The figure in brackets attached to each contour represents the depth.

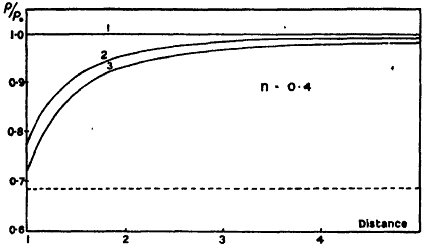


FIG. 8.—Abscissæ represent distance from centre along y-axis expressed as multiple of radius. The broken line represents density corresponding with speed of flow equal to local speed of sound.

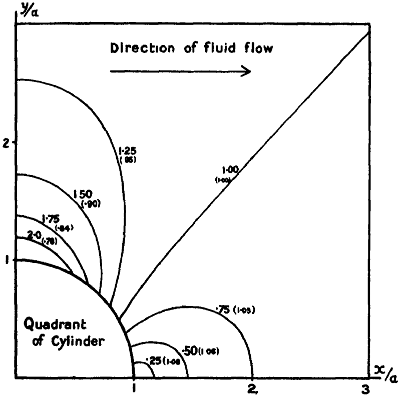


FIG. 9.—Contours of equal velocity of flow and equal depth of electrolyte, when $n = 0 \cdot 4$. The figures in brackets are the depths, a is radius of cylinders.

Cylinder moving at 0·6 times Velocity of Sound.

The next problem attempted was to find the flow past a circular cylinder placed in a stream of air moving at 0·6 times the velocity of sound. Starting again with the values of 0 characteristic of an incompressible fluid, a new bottom was found by means of the curve of fig. 6. One section of this, namely, the section by a plane through the centre of the cylinder perpendicular to the main direction of the stream is shown in fig. 10 as curve 2. It will be seen that, as might be expected, it is similar in general character to curve 2 of fig. 8 for the case where $n = 0 \cdot 4$, but that the variations in depth are greater. The measurements made with this bottom were then used to calculate a new bottom whose transverse section is shown in curve 3.

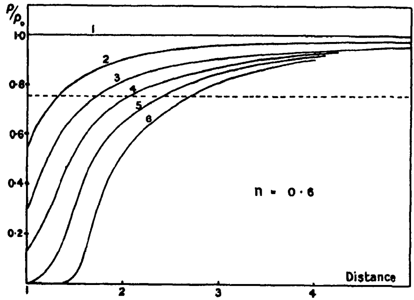


Fig. 10.—Abscissæ represent distance from centre along y-axis expressed as multiple of radius. The broken line represents density corresponding with speed of flow equal to local speed of sound.

The measurements made with this bottom were used to find a fourth bottom shown on curve 4. It will be seen that curve 4 is further from curve 3 than the corresponding curve in fig. 8.

The process was continued, but as will be seen by inspecting fig. 10 or Table 2 where the results for points on the transverse section of the field are given, there was no tendency for the process to produce a convergent solution of the problem. The shallow patch on each side of the cylinder grew progressively shallower and larger in area at an increasing rate till finally in curve 5 the depth close to the cylinder became zero. It seems clear that however far the process were continued a converging result would never be reached. Considerations of all possible alterations to the shape of the bottom, consistent with the condition that the depth at the front and back of the cylinder must correspond with zero velocity, make it appear that *there is no possible shape to which the bottom could be cut so as to represent irrotational flow of air past a circular cylinder* This amounts to the statement that no solution involving irrotational motion exists, for if such a solution could exist it would involve a definite value of ρ/ρ_0 for every point in the fluid, and if the bottom of the electric tank were cut so that the depth of the electrolyte was proportional to this, then the electric analogy would necessarily be applicable.

At first sight it seems strange that there should be no solution, but a consideration of Osborne Reynolds' equations for the flow of gas through an orifice shows that in one case at any rate we are already familiar with the phenomenon. Consider the flow of a gas through a uniform pipe of cross sectional area A_0 and suppose that at some point in its cross section it contracts to a minimum section A_1. The maximum flux of air through a pipe is determined by the condition that at the section A_1 the velocity of the stream is equal to the "local velocity of sound," *i e.*, the velocity of sound in air at the temperature and pressure which it has at that point. When this condition is satisfied the velocity of flow in the part of the pipe where the cross section is A_0 is

$$a_0 \left\{ \tfrac{1}{2} (\gamma + 1) \left(\frac{A_0}{A_1} \right)^{\frac{2(\gamma-1)}{\gamma+1}} - \tfrac{1}{2} (\gamma - 1) \right\}^{-\frac{1}{2}} = \kappa a_0, \tag{11}$$

where a_0 is the velocity of sound in air at the temperature and pressure of the air at section A_0, and κ is a fraction less than unity depending on the greatest amount of contraction which the flow experiences in its course.

It appears, therefore, that there is no solution of the equations of flow corresponding with a greater velocity at the section A_0 than κa_0 (except solutions in which the speed is *everywhere* greater than that of sound).

A particular case of a channel which contracts locally is that of a circular cylinder between two parallel walls. In this case there is therefore a maximum speed of the undisturbed stream which corresponds with a condition in which the whole of the air at the minimum section is moving with the local speed of sound. It is clear, however, that the condition that the air at all points of the minimum section should move at the local speed of sound cannot be satisfied, for curvature of the stream lines necessitates a pressure near the cylinder which is less than that near the walls. Hence the maximum possible speed of flow in the channel must be less than the maximum determined by assuming that the speed of flow at the minimum section is everywhere the local speed of sound and using formula (11). The failure of the attempt to find the flow past a circular cylinder immersed in a stream moving with a speed of 0·6 times the velocity of sound shows that in this case the necessity for having a cross-wind pressure gradient in order to make the stream lines curve round the cylinder has had a very powerful effect in reducing the maximum possible speed of the air stream.

Possibility that Maximum Possible Speed of Stream is when Local Speed of Sound is first reached at some point in the Field.

It has been seen that in the case of a local contraction in a parallel channel the maximum speed in the parallel part is attained when the speed at the point of minimum section reaches the local speed of sound. It seems worth while therefore to see whether the attainment of the speed of sound at some point in the field synchronises with the failure of convergency in the solution by means of the electrical tank. For this reason the method was applied in the case when $n = 0{\cdot}5$.

Case when $n = 0{\cdot}5$.

The values of ρ/ρ_0 in the successive approximations at points on the cross-wind section through the centre of the cylinder are given in fig. 11. It will be seen that the method does not converge but that it diverges much more slowly than it did in the case when $n = 0{\cdot}6$. It appears from the results obtained for $n = 0{\cdot}4$, $0{\cdot}5$ and $0{\cdot}6$ that the maximum speed at which there is an irrotational solution of the flow past a circular cylinder is a little less than $v = 0{\cdot}5a$. To find the value θ_s of θ at any point at which the local velocity of sound is attained we can use the formula

$$\theta_s^2 = \frac{1 + \frac{1}{2}(\gamma - 1)\, n^2}{\frac{1}{2} n^2 (\gamma + 1)}, \tag{12}$$

and the corresponding values of ρ/ρ_0 are given by substituting for θ in (10).

In this way it is found that for

$n = 0{\cdot}4,\quad \theta_s = 2{\cdot}3,\quad \rho/\rho_0 = 0{\cdot}69.$
$n = 0{\cdot}5,\quad \theta_s = 1{\cdot}86,\quad \rho/\rho_0 = 0{\cdot}72.$
$n = 0{\cdot}6,\quad \theta_s = 1{\cdot}58,\quad \rho/\rho_0 = 0{\cdot}76.$

These values of ρ/ρ_0 are marked on figs. 8, 10 and 11 as dotted lines.

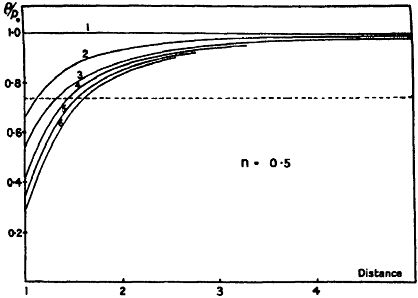


FIG. 11.—Abscissæ represent distance from centre along y-axis expressed as multiple of radius. The broken line represents density corresponding with speed of flow equal to local speed of sound.

It will be seen that for $n = 0 \cdot 4$, this line is below the lowest point in any of the curves so that the local speed of sound is never attained. For $n = 0 \cdot 6$ there is a large region in all the successive approximations where the speed is higher than the local speed of sound.

For $n = 0 \cdot 5$ in the first approximation there is a small region near the cylinder where the local speed of sound is attained, and in succeeding approximations this region increases in size.

These results therefore are consistent with the hypothesis that no irrotational flow past a circular cylinder is possible which involves the existence of a speed higher than the local speed of sound at any point on the field; but there appears at present to be no theoretical ground for supposing that this hypothesis is true in general. It is hoped that more light will be thrown on the question when cylinders of non-circular section are tested by means of the electric analogy.

Comparison of Results with Rayleigh's Approximation.

It has been pointed out that the method here described should give results which ought to be nearly, though not quite, the same as those obtained by successive applications of Rayleigh's method. The first approximation to the effect of compressibility was worked out by Rayleigh and we have compared values obtained from his formulae with those derived using the tank. In Table 2 are given the observed values of θ corresponding with successive bottoms to the tank for points at various distances y from the centre of the cylinder along a line perpendicular to the direction of the stream. In this table the column headed θ_1 is the same for all speeds of flow, and represents the values corresponding with uniform depth. θ_2 represents the first approxi-

Table II.—Table showing that Rayleigh's Approximation is very close to the Approximation θ_2 obtained by the Tank.

$n = 0{\cdot}4$.

y/a	θ_1	θ_2	θ_3	Rayleigh's approximation.
1·04	1·95	2·13	2·13	2·10
1·20	1·69	1·89	1·89	1·84
1·50	1·45	1·60	1·60	1·55
2·0	1·25	1·36	1·35	1·31
3·0	1·11	1·19	1·18	1·14

$n = 0{\cdot}5$.

y/a	θ_1	θ_2	θ_3	θ_4	θ_5	Rayleigh's approximation.
1·04	1·95	2·2	2·5	2·65	2·76	2·22
1·20	1·09	1·97	2·16	2·31	2·44	1·95
1·50	1·45	1·66	1·75	1·86	1·94	1·61
2·0	1·25	1·39	1·45	1·50	1·54	1·35
3·0	1·11	1·19	1·20	1·23	1·25	1·16

$n = 0{\cdot}6$.

y/a	θ_1	θ_2	θ_3	θ_4	θ_5	Rayleigh's approximation.
1·04	1·95	2·40	2·85			2·32
1·20	1·69	2·11	2·50	3·16		2·03
1·50	1·45	1·73	2·01	2·45	3·34	1·69
2·0	1·25	1·44	1·61	1·79	2·11	1·40
3·0	1·11	1·22	1·30	1·36	1·46	1·18

mation to the effect of compressibility, and the figures in this column should therefore be nearly the same as those calculated from Rayleigh's formulæ. These latter are given in the last column in each section.

It will be seen that Rayleigh's formulæ do in fact give results very close to θ_2. The small difference is, no doubt, due to the fact previously pointed out that the values of $\partial\phi/\partial x$ and $\partial\phi/\partial y$ which enter into Rayleigh's formulæ are taken from the previous approximation, whereas in the tank method the corresponding values are among those determined by the tank.

It seems nearly certain that if Rayleigh's method could be applied successively in the manner which he suggested, successive values of θ close to those given in the columns headed θ_3, θ_4 .. would be obtained. If this is the case Rayleigh's solution would become divergent not, as he suggested, when the speed of the stream approaches that of sound, but at a speed rather less than half of this. There is nothing in Rayleigh's analysis to indicate this possibility, and it would be very interesting if the matter could be tested by some analytical method. On the other hand, it seems likely that, as Rayleigh presumed, a failure of convergency would indicate that irrotational flow is no longer possible.

We wish to express our thanks to Sir Ernest Rutherford for permission to construct the tank and carry out our experiments in the Cavendish Laboratory, and to the Air Ministry for a grant which has enabled one of us to devote a year to this work.

The Soft X-Ray Levels of Iron, Cobalt, Nickel and Copper.

By O. W. Richardson, F.R.S., Yarrow Research Professor of the Royal Society, and F. C. Chalklin, Ph.D.

(Received July 18, 1928.)

§ 1. In a previous paper* the writers showed that a considerable number of discontinuities which had been detected in the excitation of the soft X-rays from these elements by the photo-electric method could be arranged on the assumption that they were due to transitions from two levels not far from the M_I and $M_{II\ III}$, as these had been deduced from hard X-ray emission data, and a series of virtual levels falling, with a close approximation into a Rydberg series. The present paper is an attempt to arrange the remaining discontinuities which have been observed into the same scheme. Although the identity of a level formerly referred to as A with the hard X-ray M_I level and of the higher level with $M_{II\ III}$ seems probable ultimately, it is not certain, and we do not wish to be dogmatic about it, so that for the present we shall refer to these levels provisionally as X_1 and X_2 respectively.

§ 2. *Iron* (26).

In the first place, if we consider the iron data, there are a series of discontinuities lying about 80 volts below the $X_1 - b/n^2$ values. This is shown in Table I, where the first row gives the quantum number of the "b" term, the second gives the voltage of the transition $X_1 - b/n^2$ using the values in the former paper, viz., $X_1 = 180 \cdot 0$ and $b = 2420$. The third gives the value of the experimentally observed break which we now attribute to transitions from the new level indicated as X_3 to the b/n^2 terms. The fourth row gives the difference of the second and third and is therefore equal to the difference between the calculated X_1 and the experimental X_3, on the assumption that the b/n^2 terms are the same for both transitions. There is some suggestion of a systematic deviation in these differences which may well mean that the b/n^2 levels have values slightly different when an electron is missing from the X_1 level, from those which they possess when an electron is missing from the X_3 level, In the fifth, sixth and seventh rows of this Table are given respectively the experimental values of $X_1 - b/n^2$, $X_3 - b/n^2$ and $X_1 - X_3$. These differences are more irregular than those in the fourth row.

---

* 'Roy. Soc. Proc.,' A, vol. 119, p. 60 (1928).

In Table I and similar tables in this paper, values which are used twice are indicated by a bar above, and values involving terms calculated from the Rydberg formula are enclosed in brackets. It is satisfactory that the only members of the new series which fail are those for $n = 9$ and $n = 11$, which would be expected to be weak and also are so close to the neighbouring inflections that they might be difficult to resolve.

If we disregard the small variation in the b/n^2 terms, the mean value of $X_1 - X_2$ is 79·6, giving for the value of the X_2 level 100·4 volts. The differences between the X_1, X_2 and X_3 levels are similar to, but not identical with, those of the M_I, $M_{II\,III}$ and $M_{IV\,V}$ hard X-ray levels, although they run about 80 volts deeper in each case. Having found a series of discontinuities proceeding from each of these three levels, we might from analogy with the accepted interpretation of the X-ray emission spectra have expected to have disposed of the whole "spectrum" in this range. This, however, is not the case. Thomas gives thirteen values between 180 and 331 volts for iron, and the most striking feature of our curves for iron was the very strong inflection at 288·3 volts and the absence of further discontinuities from this point up to 498 volts. This suggested that the 288·3 effect might represent the limit of a series. Returning to Thomas's values we are able to find a series of critical potentials having an almost constant difference from those of the X_1 series We shall denote this series as $X_0 - b/n^2$, X_0 denoting a level deeper than X_1.

The second row of Table II contains the calculated values of $X_1 - b/n^2$. The third row contains the experimental breaks $X_0 - b/n^2$. The fourth gives the differences of the third and second series, which, if b/n^2 is constant for corresponding terms of each series, should give the value of $X_0 - X_1$ (calc.). The fifth, sixth and seventh rows give corresponding observed data.

In this series the only missing number up to $n = 11$ is $n = 10$, and there is a very definite indication of a systematic variation of "b" with "n" when the X_0 is compared with the X_1 series. When this is allowed for the agreement is excellent, and the experimental value of the X_0 level would appear to be about three or four volts too low. This series, therefore, implies the existence of an additional level X_0 at about 292 volts, the value determined by direct experiment being 288·1 (R. & C.), 288·6 (Thomas). The variation of $X_0 - X_1$ indicates that the b/n^2 terms are a little deeper for low "n" when the electron is out of the X_0 level than when it is out of the X_1 level. We have, now, four initial states from which jumps can be made to our "b" series of terms. It seems significant that there are just four discontinuities in our soft X-ray curves in the range 450 to 750 volts for iron. This suggests that these may be

Table I.

n	4.	5	6.	7.	8	9.	10	11.	12
$X_1 - b/n^2$	(28·8)	(83·2)	(112·7)	(130·6)	(142·2)	(150·1)	(155·8)	(160)	(163·2)
Experimental break, $X_3 - b/n^2$	—	—	34·3 (2·0R)	51·3 (1·6R)	62·0 (2·7R)	67·6 ? (1·6R)	75·3 (2·2R)	—	$\overline{82\cdot7}$ (2·3R)
X_1 (cal.) $- X_3$	—	—	(78·4)	(79·3)	(80·2)	(82·7)?	(80·5)	—	(80·5)
$X_1 - b/n^2$	28·8	$\overline{82\cdot7}$	$\overline{112\cdot2}$	$\overline{131\cdot8}$	140·7	150·2	(155·8)	159·8	163·2
$X_3 - b/n^2$ $X_1 - X_3$	— —	— —	34·3 77·9	51·3 80·5	62·0 78·7	67·6 ? 83·6 *	75·3 (80·5)	(80·2) (79·6)	$\overline{82\cdot7}$ 80·5
	—	—	− 1·7	+ 0·9	− 0·9	+ 4·0 *	+ 0·9	+ 0·0	+ 0·9

Table II.

n.	4.	5.	6	7	8	9.	10	11	12.	∞
$X_1 - b/n^2$	(28·8)	(83·2)	(112·7)	(130·6)	(142·2)	(150·1)	(155·8)	(160)	(163·2)	(180)
$X_\infty - b/n^2$	$\overline{136\cdot0}$ (1·3R)	192·0 (1·8R)	221·3 (1·1R)	241·0 (1·0R)	250·8 (1·4R)	(261·5) (1·1R)	— —	271·5 (1·2R)	— —	288·6 (1·1R)
$X_\infty - X_1$ (calc.)	(107·2)	(108·8)	(108·6)	(110·4)	(108·6)	(111·5)	—	(111·5)	—	108·6
$X_\infty - b/n^2$	$\overline{136\cdot0}$	192·0	221·3	241·0	250·8	261·5	—	271·5	—	288·6
$X_1 - b/n^2$	28·8	$\overline{82\cdot7}$	$\overline{112\cdot2}$	$\overline{131\cdot8}$	140·7	150·2	(155·8)	159·8	163·2	180·0
$X_\infty - X_1$	107·2	109·3	109·1	109·3	110·1	111·3	—	111·7	—	108·6

transitions from an " L " level to the " X " levels, in which case we shall have this equation $(L - X_n) + X_n = L$, $n = 0, 1, 2, 3$. The result of applying this combination, which is probably only approximately correct (see p. 234 below), is shown in Table III.

Table III.

High volt break (R & C).	Interpretation.	Suggested final level	Value of final level.	Sum.
volts			volts.	volts.
498	$L - X_0$	X_0	288·4	786·4
608·1	$L - X_1$	X_1	180	788·1
640	$L - X_2$	X_2	151·8	791·8
689·7	$L - X_3$	X_3	100·4	790·1

In the fourth column the direct experimental value of X_0 and the average value of X_2 and X_3 have been used, and it will be seen from the last column that the agreement with the combination principle is excellent. The agreement is still further improved if, instead of $X_0 = 288 \cdot 4$, we take the value 292 volts which the numbers in Table II indicate as more probable for this level. This would give 790 instead of 786·4 in the last column. The variation of b and n, indicated in Table I for the X_3 level as compared with the X_1 level suggests that a better value for X_3 than 100·4 would be 99·5, which would give 789·2 for the last figure in the fifth column In any event the agreement is well within the experimental error. If instead of our values we use Thomas's values in this region we find that he does not record an inflection near 498 and 608·1 volts, which we have attributed to $L - X_0$ and $L - X_1$ transitions, but he gives discontinuities 639 and 704·3 volts From these we get

$$(L - X_2) + X_2 = 639 + 151 \cdot 8 = 790 \cdot 8 \text{ volts.}$$
$$(L - X_3) + X_3 = 704 \cdot 3 + 99 \cdot 5 = 803 \cdot 8 \text{ volts.}$$

He also gets an inflection at 818·5 which may indicate an L level or levels. If we take the average of all these seven values we get 796 volts, but the first five, which are almost identical, give 790·0 with a maximum deviation of ± 1·9.

These L values are much higher than the values of the L_{II} and L_{III} levels which have been obtained from X-ray emission data. Thus Thoraeus* gives for L_{II} 723, L_{III} 709 volts. He also gives $M_{II\,III}$ and measures the $L_{\beta 3,\,4}$ line from which $L_I = 818$ volts. Thus our L values are not very different from the

---

* 'Phil. Mag.,' vol. 2, p. 1007 (1926).

L_I level deduced in this way, and, in fact, one inflection measured by Thomas is in exact agreement with it. However, we do not feel that any confidence can be placed in this identification; it seems to us more likely that all our levels are deeper by an amount of the order of 80 volts than the corresponding measured values of the corresponding X-ray emission levels, as these have been interpreted up to the present. The values of these emission levels as given from experiments by Thoraeus involve the experimental determination of the limit of the K absorption edge, and in a recent paper Voorhees and Lindsay* announce that the K absorption edge of iron has a complicated fine structure extending for more than 200 volts. Accordingly it would seem that none of these levels can be determined from X-ray emission data except to within a very uncertain additive constant.

As a matter of fact our transitions give values which, while not identical with, are close to, those of what would appear to be corresponding X-ray emission transitions. Thus Thoraeus gives the following lines:—

$$\begin{aligned} L_\alpha &= L_{III} - M_{IV\,V} &&= 17 \cdot 58 \text{ Å} = 702 \cdot 2 \text{ volts} \\ L_\beta &= L_{II} - M_{IV} &&= 17 \cdot 22 \text{ Å} = 716 \cdot 9 \text{ volts.} \\ L_l &= L_{III} - M_I &&= 20 \cdot 12 \text{ Å} = 613 \cdot 6 \text{ volts.} \\ L_\eta &= L_{II} - M_I &&= 19 \cdot 65 \text{ Å} = 628 \cdot 2 \text{ volts.} \end{aligned}$$

If the L_{II} and L_{III} levels are unresolved in the soft X-ray work, the mean of the first two 709·5 volts would compare with 689·7† for $L - X_3$, and the mean of the two last 620·9 volts would compare with 608·1 for $L - X_1$. In X-ray emission the transitions from L_I and L_{II} to M_{II} and M_{III} are forbidden by the selection principle. The X-ray emission line which might be expected in the neighbourhood of our 498·1 volts discontinuity does not appear to have been looked for up to the present.

Subtracting L_l from L_α we get $M_I - M_{IV\,V} = 88 \cdot 6$ volts, subtracting L_η from L_β we have $M_I - M_{IV} = 88 \cdot 7$ volts. From the limits of the X_1 series and the X_3 series we get $X_1 - X_3 = 79 \cdot 8$, thus the differences of our new levels, while not the same, are not very far from the differences of the corresponding X-ray emission levels.

There is also some evidence of the existence of transitions within the X levels. These are collected in Table IV.

---

* 'Phys. Rev.,' vol. 31, p. 306 (1928).

† Thibaud, 'J. Physique,' vol. 8, p. 447 (1927), gives Lα as 697 volts, using a grating method.

Table IV.

Suggested transition.	Calculated critical potential	R. and C.	Thomas.
	volts.	volts.	volts.
$X_0 - X_1$	112	115·1	$\overline{112\cdot2}$
$X_0 - X_2$	140·2	138·5	$\overline{140\cdot7}$
$X_0 - X_3$	192·5	189·6	$\overline{192\cdot0}$
$X_1 - X_3$	80·5		$\overline{82\cdot7}$
$X_2 - X_3$	52·3	50·6	$\overline{51\cdot3}$
$X_1 - X_2$	28·2	(below range)	$\overline{28\cdot8}$

Although the agreement here is quite good, the fact that all the inflections have already been used as series inflections makes it less convincing than would otherwise be the case.

The arrangement of these levels, excluding the series levels, is exhibited diagrammatically in fig. 1. This arrangement accounts for the iron discontinuities found by Thomas at 28·8 and 34·3 volts, and practically all of those above 48 volts. The exceptions are:—94·8 (1·3R), 91·2 ? (1·3R), 87·0 ? (1·1R); these are all just a little less than the X_3 level and are almost certainly transitions from the X_3 level to some very superficial levels; 169·4 (2·3R), which is similarly situated just above the X_1 level and probably has a similar relation to that level; 181·6 (1·2R) which can be interpreted as the X_1 level;

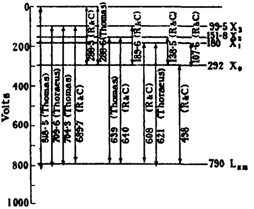


FIG. 1.—Energy Diagram for Iron (virtual orbits omitted)

212·0 (1·5R), which we were unable to find; 228·1 (2·8R) and 303·7, as to the reality of which we expressed doubt in our first paper, and 277·4 (1·4R) and 331·0 (1·6R), which also we did not find and were inclined to attribute to the presence of impurities It appears, therefore, that all the well-established inflections above 48 volts, as well as two of those below, are accounted for by the scheme of levels put forward.

§ 3. *Nickel* (28).

In the former paper, using Thomas's results, we obtained two series for nickel similar to those for iron. As in the case of iron we attribute these to transitions from provisional X_1 and X_2 levels, and a series of Rydberg terms which are almost identical in the two cases.

Table V shows that in nickel, as in iron, there is also a series of transitions from a higher X_3 level to the b/n^2 level. The second row contains the calculated values of $X_1 - b/n^2$, the third row the observed nickel inflections for $X_3 - b/n^2$, and the fourth row the difference of these quantities, X_1 calculated and experimental X_3. Rows five, six and seven contain the corresponding observed data and differences. These differences do not seem to show much evidence of systematic variation with n, the average value of $X_1 - X_3$ is 79·1, giving $X_3 = 102 \cdot 8$ volts.

We have now found nickel levels corresponding to the X_1, X_2, X_3 levels in iron. It is natural to inquire whether there is not also an X_0 level. We approached this by considering the high voltage discontinuities 774·0 and 833·4 given by Thomas for nickel, which seem to correspond to the values 639·0 and 704·3 for iron which were assigned to $L - X_2$ and $L - X_3$ respectively. We have just found that X_3 for nickel is 102·8, which in combination with 833·4 gives a calculated L level at 936·2. Using 774·0 as $L - X_2$ in combination with $X_2 = 152 \cdot 5$, we get a calculated L level at 926·5. Actually Thomas finds a discontinuity at 948·0, which may represent this L level. Taking the mean of these three values we get L = 937 volts. With nickel we found a discontinuity at 532 volts which appeared to correspond to the 498-volt discontinuity for iron, which we have attributed to the $L - X_0$ switch. On this basis the X_0 level for nickel should be at 405 volts; using this as a guide we were able to find what appears to be the $X_0 - b/n^2$ series for nickel. We have, however, only the first three terms, so we cannot regard this as being very certain. Thomas gives no value for nickel between 341·3 and 774 volts; we, however, found a discontinuity for nickel at 392·4 volts, which, although at one time suspected of being due to tungsten, appears in the light of recent

Table V.

n.	4	5.	6.	7.	8.	9.	10.	11.	12.	∞.
$X_1 - \rightarrow n$	(34·6)	(87·6)	(116·4)	(133·8)	(145·1)	(152·8)	(158·3)	(162·5)	(165·6)	(181·9)
$X_3 - b/n^2$	— —	8·6 (2·4R)	$\overline{34\cdot6}$ (2·2R)	54·8 (0·9R)	65·6 (2·0R)	71·9 (1·5R)	79·8 (1·7R)	— —	$\overline{87\cdot2}$ (1·6R)	— --
X_1 (cal.) $- X_3$	—	(79·0)	(81·8)	(79·0)	(79·5)	(80·9)	(78·5)	—	(78·4)	-
$X_1 - b/n^2$	34·6	$\overline{87\cdot2}$	116·6	133·7	145·0	(152·8)	(158·3)	(162·5)	(165·6)	-
$X_2 - b/n^2$	—	8·6	39·8? (1·4R)	54·8	65·6	71·9	79·8	(83·2)	$\overline{87\cdot2}$	
$X_1 - X_3$	—	78·6	76·8	78·9	79·4	(80·9)	(78·5)	(79·3)	(78·4)	—

evidence to be due in fact to nickel. This inflection, which was obtained with low resolving power, may well represent several transitions higher up in the series which were not resolved.

Table VI.

n	4.	5	6.	7.	∞.
$X_1 - b/n^2$	(34·6)	(87·6)	(116·4)	(133·8)	—
$X_0 - b/n^2$	257·0	311·5	341·3	—	—
Difference	(222·4)	(223·9)	(224·9)	—	181·9
$X_1 - b/n^2$	$3\overline{4}\cdot 6$	$\overline{87}\cdot 2$	$1\overline{16\cdot 6}$		—
$X_0 - b/n^2$	257·0	311·5	341·3	—	409
$X_0 - X_1$	222·4	224·3	224·7	—	227

The data for testing the $X_0 - b/n^2$ level are set out in the usual form in Table VI. It will be noticed that there appears to be a systematic deviation in the differences with different values of n of a similar character to that found for iron. This indicates a value for the difference as n approaches ∞ of about 227, which would give a value for the X_0 level of about 409 in satisfactory agreement with the value got from the 532 volts inflection in combination with the L and the other X terms.

The comparison of these values with those got from X-ray emission data gives similar results to those found with iron. Taking the mean of Thoraeus's value for the L_α and L_β lines as representing the transition $L_{II\,III} - M_{IV\,V}$ we get 858·5 volts, the inflection obtained by Thomas and interpreted as $L - X_3 = 833\cdot 4$ volts. The mean of Thoraeus's L_l and L_η values gives the mean for the transition from $L_{II\,III} - M_I$ as 754·6 volts. Thomas does not find this discontinuity. If, however, we use the value of $X_1 - X_3$ from our critical potential data and subtract it from Thomas's inflection, which we interpreted as $L - X_3$, we get a result 754·3 volts; similarly, subtracting $X_1 - X_3$ from the 774 inflection attributed to $L - X_3$ we get for $L - X_1$ 744·7 volts, for $M_I - M_{IV\,V}$ we get from the X-ray emission data 103·9 volts as compared with 79·1 for $X_1 - X_3$ from the soft X-ray emission levels. Whilst these differences are not great, it is thought that they exceed the probable experimental errors in some cases.

On the other hand, if instead of dealing with the transition differences corresponding to lines, we consider the absolute values of the levels, we find for nickel, just as for iron, a much greater divergence between the values got from the soft X-ray and the X-ray emission data. Thus the mean of Thoraeus's values for the $L_{II\,III}$ levels for nickel is 867 volts as compared with the average value from soft X-ray data of 937 volts and an inflection which has been interpreted as a soft X-ray determination of this level as 948 volts. This discrepancy is within a few volts of a similar one got with iron. The evidence bearing on the possible existence of inflections corresponding to transitions within the X levels is given in Table VII.

Table VII.

	X_0-X_1	X_0-X_2	X_0-X_3	X_1-X_2	X_1-X_3	X_2-X_3
Calculated	227·1	256·4	306·2	29·3	79·1	49·8
Experimental break	228·2	$\overline{257\cdot0}$	abs	31·0	79·8	51·3
Character and allocations	(1·1R)	(1·6R)X_0(4)	,,	(1·4R)	(1·7R)X_2(10)	(2·0R)

The calculated $X_0 - X_3$ transition at 306·2 may possibly be mixed up with 311·5 (1·8R) which has been assigned to X_0 (5).

If, as in dealing with similar inflections in the case of iron, we attribute the following, namely, 98·3 (1·7R), 94·9 ? (1·2R), 91·2 ? (0·9R), to a short series of transitions from the X_3 level to some very superficial levels, and 139·6 (1·4R) to a similar transition from the X_2 level, and also 178·2 (1·2I) to a similar transition from the X_1 level, all the nickel inflections recorded by Thomas between 30 and 180 volts are accounted for, except 44·7 (1·5I). It leaves, however, nine inflections between 180 and 300 volts which, up to the present, we have been unable to account for. We only found three of these, and it is possible some of them may be due to contamination.

§ 4. *Cobalt* (27).

It is unnecessary to describe in detail the series for this element. They correspond to those of iron and nickel. In Table VIII we give the $X_3 - b/n^2$ discontinuities below the calculated $X_1 - b/n^2$ values shown in the second row and below the observed $X_1 - b/n^2$ discontinuities in the fifth row The corresponding differences are shown in the fourth and seventh rows respectively. With this element there seems to be an indication of a slight systematic variation in the opposite direction to that found with iron, the differences being larger for the smaller values of n. If we take 79 as an estimate of the

Table VIII.

n.	4.	5.	6	7	8.	9.	10.	11.	12.	∞.
$X_1 - b/n^2$ (calculated)	(32·2)	(85·2)	(114·0)	(131·4)	(142·7)	(150·4)	(155·9)	(160·1)	(163·2)	179·5
$X_2 - b/n^2$	—	—	$\overline{32\cdot2}$ (2·2 R)	49·3 (1·8 R)	62·8 (1·7 R)	70·7 (1·1 R)	76·7 (1·8 R)	81·0 (1·2 R)	$\overline{84\cdot3}$ (1·9 R)	—
$X_1 - X_2$	—	—	(81·8)	(82·1)	(79·9)	(79·7)	(79·2)	(79·1)	(78·9)	—
$X_1 - b/n^2$	$\overline{32\cdot2}$	$\overline{84\cdot3}$	$\overline{114\cdot3}$	$\overline{131\cdot4}$	143·8	(149·7)	154·0	(160·0)	163·5	—
$X_2 - b/n^2$	—	—	$\overline{32\cdot2}$	49·3	62·8	70·7	76·7?	81·0	$\overline{84\cdot3}$	—
$X_1 - X_2$	—	—	82·1	82·1	81·0	(79·0)	77·3	(79·0)	79·2	—

value of $X_1 - X_3$, when *n* is large, we get 100·5 volts as the value of the X_3 level.

We have not examined this metal, and have only Thomas's data to work upon. It seems clear that the inflection at 705·2 volts corresponds to the inflections which we have assigned to $L - X_2$ for iron and nickel, and the one at 764·8 volts to those which we have assigned to $L - X_3$ for iron and nickel. There is no recorded inflection which corresponds to the $L - X_1$ switch. The combination $(L - X_2) + X_2 = 705·2 + 150·5$ gives $L = 855·7$ volts, whilst $(L - X_3) + X_3 = 764·8 + 100·5$ gives $L = 865·3$ volts. Thomas's inflection for this is evidently the one at 873·2 volts. The mean value of these three L level estimates is 864·4 volts.

There is no evidence of the $L - X_0$ transition in Thomas's data, but it seemed reasonable to expect that the X_0 level of cobalt would lie somewhere between the corresponding levels for iron and nickel. Working on this assumption, we found the series shown in Table IX, in which the second line gives the calculated values of $X_1 - b/n^2$, the third and sixth rows the observed values of $X_0 - b/n^2$, the fifth row the observed values of $X_1 - b/n^2$, and the fourth and seventh rows the corresponding differences. In each case these differences show a systematic deviation of the same character as that observed for the corresponding differences for iron and nickel.

Table IX.

n.	4.	5.	6.	7.	8.	∞
$X_1 - b/n^2$ (calc.)	(32·2)	(85·2)	(114)	(131·4)	(142·0)	--
Observed breaks	223	277·4	306·5	326·3	335·8	--
$X_0 - b/n^2$	(1·4R)	(1·8R)	(2·1R)	(1·2R)	(2·3R)	
$X_0 - X_1$	(190·8)	(192·2)	(192·5)	(194·9)	(193·8)	—
$X_1 - b/n^2$	32·2	84·3	114·3	131·4	143·8	179·5
$X_0 - b/n^2$	223·0	277·4	306·5	326·3	335·8	375
$X_0 - X_1$	190·8	193·1	192·2	194·0	192·0	195·5

The value of $X_0 - X_1$ for large *n* appeared to be about 195·5, which gives a value of X_0 about 375 volts.

For cobalt Thoraeus's measurements of the X-ray lines $L_\alpha - L_\beta$ give $L_{II\,III} - M_{IV\,V} = 782·3$ volts. This compares with Thomas's inflection at 764·8 volts assigned to $L - X_3$. Thoraeus's values for $L_l - L_\eta$ give for $L_{I\,III} - M'_I$ 686·2 volts. From $L - X_3 = 764·8$ and $X_1 - X_3 = 79$ we get $L - X_1 = 685·8$. From $L - X_2 = 705·2$ and $X_2 - X_1 = 29·0$ we get $L - X_1 = 686·2$: both these values are identical with each other and with

$L_{II\,III} - M_I$ to within the limits of experimental error. The value of $M_I - M_{IV\,V}$ from Thoraeus's data is equal to 96·1 volts, whereas these critical potential data give for $X_1 - X_2$ 79 volts; the value of the L level got from the X-ray emission data of Thoraeus's 791 volts, compared with the mean of the two combination soft X-ray estimates of 860 volts and the direct soft X-ray determination of 873·2 volts. These discrepancies are precisely similar to those found with iron and nickel.

The evidence for the occurrence of transitions between the X levels are given in Table X.

Table X.

	$X_0 - X_1$.	$X_0 - X_2$.	$X_0 - X_3$	$X_1 - X_2$.	$X_1 - X_3$.	$X_2 - X_3$.
Calculated	195·5	224 5	275 6	29 0	80 1	51·1
Experimental break	195·6	$\overline{223}\;\overline{0}$	$277\;\overline{4}$	29·3	$81\;\overline{0}$	52·8
Character and allocations	(1·6R)	(1·4R)X_0(4)	(1 8R)X_0(5)	(1·2R)	(1·2R)X_2(11)	(1·2R)

If, as in dealing with the similar inflections in the case of iron and nickel, we assign the following, viz., 97·7 (1·7R), 92·5 ? (1·5R), 89·4 ? (1·3R), which are a little less than the X_3 levels to a short series of transitions from the X_3 to a set of superficial levels, and 138·3 (1·5R) to a transition from the X_2 level to a superficial level, and 185·5 (1·2R) to the X_1 level itself, these schemes account for all Thomas's inflections for cobalt above 25·7 volts, with the exception of 38·5 (1·9R), 43·2 (1·6I), 214·0 (1·3R) and six discontinuities between 223·0 and 306·5. One of these six is identical with the nickel inflection X_0 (4); the other five are practically identical with unclassified nickel inflections, one of which is identical with carbon K and two others with well-marked tungsten levels, so that it seems doubtful whether these levels which fall outside the scheme really belong to cobalt or nickel.

§ 5. *Copper* (29).

Our results with this element are based entirely on the data given in the papers by Thomas,* and Compton and Thomas.† The $X_3 - b/n^2$ series is given in Table XI (third and sixth rows). The second row gives the calculated, and the fifth row the observed values for $X_1 - b/n^2$, corresponding sets of differences

---

* 'Phys. Rev.' vol. 26, p. 739 (1925).

† 'Phys. Rev.' vol. 28, p. 601 (1926).

Table XI

n	4.	5.	6.	7.	8.	9	10.	11.	12.	x
$X_1 - b/n^2$ (calculated)	(37·5)	(90·6)	(119·4)	(136·8)	(148·2)	(155·9)	(161·4)	(165·6)	(168·7)	(185)
Observed breaks	—	10·3	$\overline{37\cdot5}$	53·8	—	74·0	—	84·0	88·0	—
$X_3 - b/n^2$	—	(2·3R)	(2·8R)	(2·2R)	—	(1·6R)	—	(1·8R)	(1·4R)	
$X_1 - X_3$	—	(80·3)	(81·9)	(83·0)	—	(81·9)	—	(81·6)	(80·7)	—
$X_1 - b/n^2$	37·5	92·2	$\overline{117\cdot8}$	(136·8)	148·4	157·1	(161·4)	(165·5)	168·3	—
$X_2 - b/n^2$	—	10·3	37·5	53·8	(66·5)	74·0	(79·8)	84·0	88·0	—
$X_1 - X_2$	—	81·9	80·3	(83·0)	(81·9)	83·1	(81·6)	(81·5)	(80·3)	—

being given in the fourth and seventh rows. The variations in these differences seem irregular: it is doubtful whether they are systematic. We have, therefore, taken the average of 81·7, which would also seem to be not far from the value for large n, as the difference of the terms $X_1 - X_3$. This gives for the X_3 term the value 103·8 volts.

Thomas's discontinuity at 929·0 volts for copper appears to correspond with the inflections assigned to the $L - X_3$ transitions for iron, cobalt and nickel, but the copper inflection at 820·4 volts seems more likely to correspond to $L - X_1$ than to the discontinuities which have been assigned to $L - X_2$ for iron, cobalt and nickel. With these identifications we have $(L - X_3) + X_3 = 929 \cdot 0 + 103 \cdot 4$ gives $L = 1032 \cdot 4$, and $(L - X_1) + X_1 = 820 \cdot 4 + 185$ gives $L = 1005 \cdot 4$. The mean of these is 1018·9, and Thomas gives a discontinuity at 1017 volts which we can therefore assign to the L limit. The mean of these three values is $L = 1018 \cdot 6$. Thoraeus's emission data give $L_{\text{II III}} - M_{\text{IV V}} = 937 \cdot 4$ volts as compared with 929·0 for $L - X_3$ and $L_{\text{II III}} - M_{\text{I}} = 822 \cdot 5$ volts as compared with 820·4 for $L - X_1$. From the X-ray data we have $M_{\text{I}} - M_{\text{IV V}} = 114 \cdot 9$ volts as compared with $X_1 - X_3 = 81 \cdot 7$. There is a discrepancy between the value of $L_{\text{II III}}$ got from X-ray emission data, and the present soft X-ray L value which is similar to that already noted with iron, cobalt and nickel.

It is unfortunate that the results of Compton and Thomas go no higher than 280 volts. If we extrapolate the X_0 level from the value for the elements iron, cobalt and nickel we get for copper $X_0 =$ about 420 volts. The first line of the series $X_0 - b/n^2$ should be at about 269 volts. There is a break at 269·6 (1·3R) (?), but the second critical potential is well beyond the range. The basis for the existence of transitions between the different X levels is exhibited in Table XII.

Table XII.

	$X_0 - X_1$	$X_0 - X_2$.	$X_0 - X_3$.	$X_1 - X_2$.	$X_1 - X_3$.	$X_2 - X_3$.
Calculated	(235)	(236·9)	(316·6)	(28·9)	(81·6)	(52·7)
Experimental break	235·6	abs.	*	abs.	84·0 ?	53·8
Character and allocations	(1·9R)	—	—	—	(1·8R)X_3(11)	(2·2R)X_3(7)

* This place is beyond Compton and Thomas's range.

The definite absence of $X_0 - X_2$ and of $X_1 - X_2$ and a bad agreement with $X_1 - X_3$ strongly suggests that these transitions do not occur with copper.

In the case of copper, a very considerable number of discontinuities have been recorded by Thomas and Compton and Thomas which do not seem to fall into this scheme.

§ 6. *General Considerations.*

The values in volts of all the L and X levels are collected in Table XIII, and the values of the square roots are plotted as a Moseley diagram against atomic number in fig. 2.

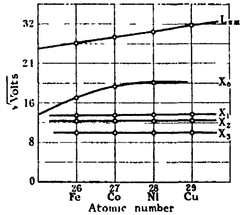


FIG. 2.—Moseley Diagram for Soft X-ray Levels.

Table XIII.

	Fe.	Co.	Ni.	Cu
L	790	860	934·7	1018·9
X_0	292	375	409	(420)
X_1	180	179·5	181·9	185
X_2	151·8	150·5	152·6	156·1
X_3	99·5	100·5	102·8	103·8
$X_1 - X_3$	80·5	79	79·1	81·7
$M_I - M_{IV\,V}$ from X-ray data	88·65	96·1	103·9	114·9
Difference	8·15	17·1	24·8	33·2

The following points may be noticed :—

(*a*) The L levels fall very accurately on a straight line, although they are plotted to represent discontinuities due to unresolved levels about 20 volts apart. The form of this curve is in agreement with the view that this level is unaffected by the external atom-building.

(*b*) The X_1, X_2 and X_3 levels are practically the same for each element. The slight variations that are shown are common to all these levels, for instance, in cobalt X_1, X_2 and X_3 are all low, and in copper they are all high.

(*c*) The X_0 level shows a variation which is very similar to the shape of the Moseley diagrams of the L and M levels when they are being completed. This fact may throw light on the peculiar constancy of the X_1, X_2 and X_3 levels. If the X_0 level were shielding the X_1, X_2 and X_3 levels, the stationary condition of these levels would be accounted for. In fact, if the X_0 level were absent, it would be difficult to account for this phenomenon. In addition, if the number of electrons present in the X_0 level increases by unity from element to element as the atomic number increases, the approximate constancy of the numerator, *b*, for all the virtual orbit levels b/n^2 would also be accounted for.

(*d*) The difference between the values of $M_I - M_{IV\,V}$ given by X-ray emission data and of $X_1 - X_3$ seem very systematic. the successive differences for Fe, Co, Ni and Cu in Table XIII being almost exactly in the ratio 1 2 : 3 : 4. The values of $M_I - M_{IV\,V}$ should not in any event be identical with $X_1 - X_3$ even if the X and M levels are the same, since the X-ray data apply to an atom in which there is an electron missing from the L level, whereas this level is filled when the values of X_1 and X_3 are being determined.

(*e*) Some of these soft X-ray level differences show a surprising agreement with corresponding X-ray emission transitions. This is especially true of the set shown in Table XIV. The first row gives the values of $L_{II\,III} - M_I$, given by the mean of Thoraeus's measurements of the L_l and L_η lines, whilst the second row gives the value of $L - X_1$ derived in various ways from the soft X-ray discontinuities. The divergence in the case of Fe might be due to the possibility that the direct soft X-ray experimental value corresponds to the transition $L_{III} - M_I$ only, for which the X-ray emission value is 613·6 volts.

It is interesting to note that in discussing our measurements in our last paper we stated that the value 608·1 might be a few volts too low.

(*f*) In Table XIV we set out in the third, fourth and fifth rows the estimations of the corresponding L levels got by three different methods The divergence between the soft X-ray combination values and the X-ray emission values shown in the last row, is nearly constant. It is possible that this divergence may disappear when the complex K X-ray absorption spectra of these elements is better understood.

The discrepancy between the values of "L by combinations" and "L direct" is likely to be real. In fact, the combination principle as it has been used in this paper for getting L is almost certainly inexact. For example, $L - X_1 = 608 \cdot 1$ means the removal of an electron from the L level to the X_1 level, so that at the end of this transaction the X_1 has one electron more

Table XIV.

	Fe	Co.
	$L_{II\,III}-M_I = 620\cdot9$ $L-X_1 = 608\cdot1$	$L_{II\,III}-M_I = 686\cdot2$ $(L-X_3)-(X_1-X_2)=685\cdot7$
L by combinations	790·0	860·0
L direct ..	818·5	873·2
$L_{II\,III}$	716	791
L com . $\Delta L_{II\,III}$	74	69
	Ni	**Cu.**
	$L_{II\,III}-M_I = 754\cdot6$ $(L-X_3)-(X_1-X_2)=754\cdot3$	$L_{II\,III}-M_I = 822\cdot5$ $L-X_1=820\cdot4$
L by combinations	934·7	1018·9
L direct	948	1017
$L_{II\,III}$ ..	867	942
L com $\Delta L_{II\,III}$.	68	77

than the normal number, and the L level one less. X_1 means the removal of one electron from the normal X_1 level to infinity. L means the removal of an electron from the normal L level to infinity, the X_1 level being in the normal state. Thus the transactions on the two sides of the equation $(L - X_1) + X_1 = L$ do not balance but differ by the difference of the shielding of the X_1 level by an electron when in the L level in the one case and in the X_1 level in the other. It is to be expected that " L direct " will be $> (L - X_1) + X_1$. This is found to be the case for Fe, Co and Ni.

(*g*) The constant *b* in the virtual orbit levels b/n^2 corresponds in each case to an effective nuclear charge of 13·2*e* to 13·3*e*. If the X levels are all given an effective quantum number about 3, which they would have if they were M levels, their effective nuclear charges are roughly 14, 11, 10 and 8 for X_0, X_1, X_2 and X_3, respectively.

(*h*) The case for the inclusion of copper in this scheme is not so secure as that for iron, cobalt and nickel. This is partly due to the fact that the data are less complete for Cu, but even when they are full the agreement is not so satisfactory and there is a large number of surplus discontinuities. In any event there are important differences between copper and the other metals. The transition $L_{II\,III} \rightarrow X_2$ is not present and the $X \rightarrow X$ transitions are absent. In fact, copper seems to show a more normal behaviour and the usual prohibitions seem to hold.

(*i*) It is not improbable that we are dealing here with the excitation of the soft X-ray spectra of the ionised elements (spark spectra) and that the X_0 levels correspond to M_V, which is filling up at this stage and which, owing to the increased electric field caused by the ionised state of the atom, has got under some of the other M levels which are represented by X_1, X_2 and X_3. It is also very likely that the solid state of the elements will modify very considerably the behaviour of some of these rather superficial levels. This point of view seems capable of accounting for the discrepancies between the values found for the X levels and the values of the M levels got from X-ray emission data. It does not, however, seem worth while to pursue this aspect of the matter further at present as the recently discovered complexity in the X-ray phenomena for these elements makes it doubtful to what extent the discrepancy is real. The set of virtual levels b/n^2 seems to be a new phenomenon in the X-ray field.

[*Note added October 9th,* 1928. -Further measurements carried out in the Wheatstone Laboratory by Miss Andrewes and by Messrs. Bandopadhyaya and Ramachandra Rao since this paper was written show that the systems of levels here described are only parts of much larger systems; so that it now seems very improbable that there is any association between the X and M levels.]

An Investigation of Some Banded Structures in Metal Crystals.

By C. F. Elam, D.Sc., with an appendix by Prof. G. I. Taylor, F.R.S.

(Communicated by G. I. Taylor, F.R.S.—Received August 28, 1928.)

[Plate 3.]

Crystals of the native metals gold, silver and copper have been found which exhibit twinning on an octahedral plane of the spinel type. Metallurgists have described as twins the banded structures which are of very frequent occurrence in these metals when prepared commercially, as they resemble lamellar twinning in calcite and felspar very closely. A few quantitative measurements have been made to determine the relationship between such bands in a metallic crystal; but the value of the results depends on the correct determination of the orientation of the respective parts, and this can only be done by means of X-rays or by indirect methods such as the measurement of slip-bands, etching-pits, etc., which are not always trustworthy.

McKeehan* measured some structures resembling twins in nearly pure iron by obtaining reflections from the deeply etched surface of a wire mounted in a goniometer. He concluded that the crystals were twins in a crystallographic sense, the twin plane being of the form {211} and the twin axis of the form [111].

A. J. Phillips† made some measurements of the angles between more than one set of bands in the same crystal in specimens of copper and brass, and found that they were inclined at approximately 70° to each other. (The angle between octahedral planes of a cubic crystal is 70° 31′.) Similarly, he measured the inclinations of traces of secondary bands in the principal bands, and found that these also could be considered as being traces of octahedral planes in the bands, provided that the two were related to each other as the two parts of the spinel twin. From this he concluded that the bands were true twins. No direct measurements of the orientation of either the original crystal or of the bands were made, and in many cases the bands could not be traced over the edge between the two faces of the specimen, so that there was a certain amount of conjecture as to whether two traces really belonged to the same plane.

In a letter to 'Nature'‡ the present writer pointed out that banded structures in aluminium and copper had been obtained which resembled twins in appearance, but which did not have the correct relationship for normal twins. Also, in a paper describing some experiments with crystals of a copper-aluminium alloy,* two examples were given of twin-like structures. One could be described as parallel growth, part of the crystal being rotated 7° with regard to the other about a common axis. The other was the result of distortion. The crystal was so oriented that the shear stress was nearly equal on two octahedral planes and one part slipped on one plane and one on the other.

---

* 'Amer. Inst. Mining Met. Eng.,' September, 1927.

† 'Amer. Inst. Mining Met. Eng.,' February, 1927.

‡ 'Nature,' vol. 120, p. 259 (1927).

Some measurements have now been made in order to obtain further information about these structures, and to determine their true nature with greater accuracy than before.

The most complete measurements have been made on some samples of aluminium that contained exceptionally large bands. Although aluminium has the same crystal structure as copper, silver and gold (face-centred cubic), the crystals are rarely founded with a banded structure after treatment which would normally produce it. When, however, a large aluminium crystal is strained about 10 per cent and heated until it recrystallises, the new crystals have very straight boundaries, and some of these take on the typical banded structure of twinned copper.

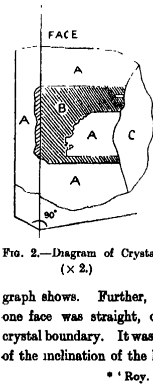


Fig. 2.—Diagram of Crystal. (× 2.)

Two examples with very clear banding were chosen and the orientation of each part relative to the surface determined by means of X-rays. The inclinations of the bands on two surfaces were also measured so that the plane of intersection between the bands and the original crystal could be compared with the X-ray measurements. Fig. 1 (Plate 3) is a photograph showing two sides at right angles, and fig. 2 a diagram of one crystal cut from a sheet $\frac{1}{8}$ inch thick. The principal crystal went through the sheet, but the twin did not, as the photograph shows. Further, although the junction between the two parts on one face was straight, on the other it was uneven and like an ordinary crystal boundary. It was not possible to obtain any trustworthy measurements of the inclination of the line of intersection on this face. Particulars of the

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 694 (1927).

X-ray data are given in Table I, and are plotted on a stereographic diagram in fig. 3. The plane of the paper represents the face of the specimen and the

Table I.—Reference Plane = Face of Specimen.

θ = Angle between normal to plane and line of intersection of face and edge.

ψ = Angle between plane containing the normal of the reflecting plane and face of specimen, measured anti-clockwise

			Reflecting Planes.		Twin Plane.
1.	A		$\theta = 115°$	$\theta = 136°$	$\theta = 157° 30'$
			$\psi = 7°$	$\psi = 67°$	$\psi = 359° 30'$
	B		$\theta = 86°$	$\theta = 123°$	$\theta = 156°$
			$\psi = 8°$	$\psi = 64°$	$\psi = 0°$
2.	A		$\theta = 83° 10'$	$\theta = 57° 30'$	$\theta = 153° 30$
			$\psi = 8° 30'$	$\psi = 81°$	$\psi = 14°$
	B		$\theta = 99° 30'$	$\theta = 42° 10'$	$\theta = 154° 30'$
			$\psi = 9°$	$\psi = 5° 30'$	$\psi = 14° 30$
		Plane of intersection of A and B,			$\theta = 148° 30'$
					$\psi = 14° 30'$

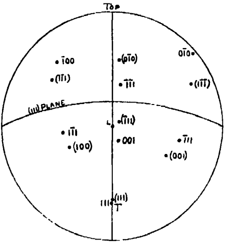


FIG. 3.—Al Twin I. Face of specimen represented by plane of paper L = Normal to Face. T = Twin Plane

vertical line corresponds to the line of intersection of the section with the face. Only the poles of those planes lying on the upper hemisphere are marked and the indices of those of crystal B are in brackets to distinguish them from crystal A. It will be seen that both crystals have an octahedral plane, 111 and (111) in the diagram marked T, in the same position. This plane has been drawn in the diagram, and it will be seen that it cuts the circumference at 90°

to the axis. This agrees with the inclination of the junction of the two parts on the face of the specimen. The centre of the projection L represents the normal to the face. In order to compare the two crystals, a projection was made of the poles of all the other planes on to the common octahedral plane. This was equivalent to rotating the figure—using a stereographic net—until the poles of the octahedral planes occupied the centre of the diagram and all the other points were moved through the same number of degrees. Fig. 4 is the

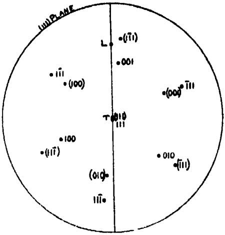


Fig. 4.—Al Twin I. Projection of both crystals on to Common Octahedral Plane.

diagram so obtained. It shows a symmetrical figure of six-fold symmetry, made up of the poles of planes belonging to the two crystals rotated 60° to each other. These crystals are therefore related in a manner similar to that of the most common type of twinning in cubic crystals, viz., rotation about a [111] axis and on a {111} plane. It is difficult to determine the plane of composition of this twin as the boundary on one section is very uneven. On this face the general direction of the plane does not agree with the trace of the twin plane, but it has been pointed out already that it does agree on the other face. The plane of composition need not be the twin plane, but it is possible that in this case it is made up of steps too small to distinguish, the majority of which agree in direction with this plane.

Another specimen was measured in the same way with the same result. The twin crystals penetrated the sheet and the boundaries were straight on both sections, so that the inclination of the planes could be determined. The results of all measurements are given in Table I, but as this crystal was so similar to the first no diagrams are given.

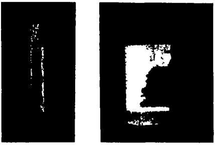


Fig 1 (Natural size)

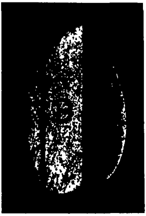


Fig. 5. (× 2)

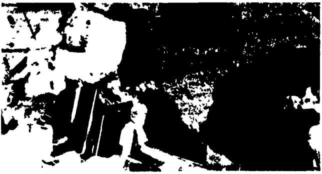


Fig. 8. (× 3)

These crystals, therefore, fulfil all the requirements for true crystallographic twinning, and it has thus been established that twins occur in aluminium, although rarely.

Another example of banded structure was also investigated, of which a photograph is given in fig. 5 (Plate 3). A crystal of aluminium in the form of a round bar pulled in tension deformed non-uniformly with the formation of a number of parallel planes running vertically along the specimen. The photograph is of a section through the test-piece. The difference of orientation of the bands is revealed by the etching and the sharp edge and re-entrant angle, both of which ran the whole length of the specimen. The section has all the appearance of lamellar twinning with nothing to distinguish it from previous examples investigated. A Laue photograph was taken, perpendicular to the plane of union, through the two parts marked A and B (fig. 5). These were very similar and were very nearly mirror images of one another. Some of the principal zones were identified and stereographic diagrams prepared (figs. 6 and 7). The zones drawn are actually those obtained from the Laue photos and bring out the symmetry of both crystals. The plane of the paper represents the plane of the section, and the vertical line the junction between crystals A and B. The centre of the diagram is the normal to the section and the axis of the bar. In both cases a (121) plane lies perpendicular to the axis.*

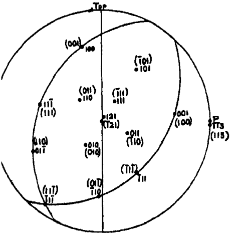


FIG. 6.—Mechanical Twin A.

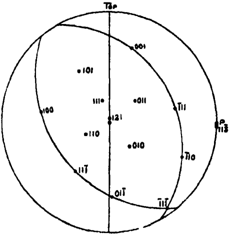


FIG. 7.—Mechanical Twin B.

---

* In an appendix Prof. Taylor has shown how these two parts can be derived from one crystal by distortion, and in order to show the relationship has given other indices to fig. 7. These are put in brackets. The unbracketed indices are what would be assigned normally to the crystals if their previous history were unknown.

It has already been shown that when an aluminium crystal is distorted in tension, the final position of the crystal axes is such that the pole of a {112} plane coincides with the axis of the specimen, and two octahedral planes, the slip planes, make equal angles with the axis. This condition has been fulfilled in both parts of the crystal, so that both are identical with regard to the vertical axis. They are also related in such a manner that the two are mirror images of each other about their plane of union, *i.e.*, the vertical plane represented by a straight line in both figures. They are also mirror images of each other in the (121) plane which they have in common. The one can be derived from the other by a rotation of 66° about an axis normal to the (121) plane. The point P on the circumference is the pole of the plane of intersection of the two crystals. The nearest crystallographic plane in both has the indices {113}, so that the specimen as a whole can be considered as a reflection twin about this plane. It is very different from the other samples already described, and the explanation must be sought in the method of production. In this respect it resembles the copper-aluminium crystal already referred to, and it is probable that here also the crystal was in such a position that slip occurred on different planes in different parts, but the original position of the axes is unknown. It is certain, however, that if this structure can be called a twin it is of a different kind from those produced during recrystallisation after mechanical and heat treatment. It is interesting to compare the similar but opposite observations with zinc. Mathewson and Phillips* showed that there is good reason to think that the banding produced by mechanical deformation of zinc is the result of twinning. Twinning in annealed zinc, however, has never been observed.

It has already been mentioned that banded structures, commonly described as twin crystals, are always found in copper, silver and gold which has been heated after being mechanically deformed. The twins so produced are generally small and frequently themselves contain twin bands. A piece of native copper was obtained which had relatively large simple twins, so that X-ray photographs could be taken of each part of the twin. There did not appear to be any relation between the parts, as was mentioned in the letter to 'Nature,' but it was subsequently found that the metal may have been distorted, so that these observations are not reliable. Later, a strip of silver was obtained $\frac{1}{16}$ inch thick, in which the crystals were large, and in some of these the twins were relatively simple and larger than usual. A photograph is shown in fig. 8, Plate 3. Even here it was extremely difficult to find examples in which the twin bands could be traced on both the face and edge of the specimens and in which they were sufficiently wide to give reliable X-ray reflections. Owing to the fact that the bands were narrow and frequently intersected by similar smaller bands, no reliable data were obtained for fixing their orientation with regard to the main crystal. This has had to be abandoned until more suitable samples can be procured. In one case the orientation of the principal crystal was determined, however, and correlated with the traces of three twin planes. A diagrammatic drawing is given in fig. 9. The traces of two planes, B and C, were clearly visible on the face and section of the specimen and on the opposite side, the crystal penetrating the strip, that of the third was only visible on the faces, but was parallel to the edge, and it was concluded that the plane was parallel with the section. A stereographic diagram was therefore made embodying the results of these and the X-ray measurements. The observations of the traces of the planes are marked with a circle. The poles of the four octahedral planes determined from X-ray photographs are marked, and the traces of three of these drawn. These two sets of measurements agree within the limits of experimental error. It is considered, therefore, that these bands follow the octahedral planes of the crystal in which they occur, but it has not yet been proved that the material within the bands bears a twinned relationship to the original crystal. It is highly probable that this is so, as Phillips concluded from his observations, but it has been shown above that apparently similar structures may or may not be considered as simple twins, and only a complete analysis of the crystal orientations can decide the question.

---

* 'Amer. Inst. Mining Met. Eng.' (1927).

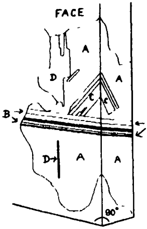


FIG. 9.—Silver Twin. (× 4.)

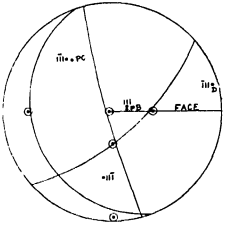


FIG. 10.

Conclusion.

It has been shown that aluminium can form twins of the spinel type, although this metal has been generally considered to be the exception amongst others of the same crystal structure to forming such twins. The plane of composition can be the twin plane, but it is not always so, as the one example illustrated shows. In this case the two components were united in one plane by a ragged boundary, indistinguishable from an ordinary crystal boundary, and this suggests the possibility that twins may be more common than appears, but that they are overlooked because they are not united along the twin plane *i.e.*, they do not exhibit straight boundaries.

There are also similar structures, *i.e.*, banding with straight boundaries, which are not of the spinel type of twin. The plane of composition in one example investigated had indices {113}, and the crystal can be considered as a reflection twin in this plane. This can be described as " mechanical twinning." Another type of banding found in the aluminium-bronze crystal referred to previously -which was only one example of many—may be more truly described as parallel growth.

This work was carried out at the Royal School of Mines, in the laboratory of Prof. H. C. H. Carpenter, F.R.S , to whom I wish to express my thanks.

APPENDIX.

By G. I. TAYLOR, F.R.S.

It is of interest to trace the connection between the known manner in which aluminium crystals are distorted under tensile stress and the formation of the " mechanical twins " discussed by Dr. Elam. As she points out, " mechanical twins " are likely to be found when two possible planes of slip are equally inclined to the axis of the specimen. There are two ways in which this condition may be satisfied. The axis of the specimen may lie (*a*) in an octahedral plane, (*b*) in a cubic plane. In case (*a*) slipping on one of the possible slip planes brings the other plane into such a position that the component of shear stress is greater on the second than it is on the first. This tends to make the material slip on the second plane and stop slipping on the first, thus giving rise to equal double slipping throughout the material.

In case (*b*), on the other hand, slip on one of the two possible slip planes increases the component of shear stress on the second plane above that on the first, so that slipping tends to proceed on whichever of the two planes it began. When slipping begins on different planes in different parts of the same specimen, it can only continue in this way provided that neighbouring parts in which the slipping is on two different planes continue to fit together as distortion proceeds. If the slipping in the two neighbouring parts is confined to one plane in each part, this condition can only be satisfied if (1) the surface of separation is a plane which contains the axis of the specimen, (2) the crystal axes on the two sides are in the position of mirror images in this plane, (3) the amount of slipping is equal in the two parts.

Dr. Elam has shown by means of the microphotograph, fig. 6, and her X-ray analysis that conditions (1) and (2) are satisfied. It seems certain that condition (3) is also satisfied, for the two parts into which the specimen was divided, both stretched through the same amount.

It remains to discuss the change in orientation of the crystal axes with respect to the axis of the specimen and the plane of union. Fig. 11 is a stereographic diagram showing the axes of a cubic crystal. A is a point on its cubic plane (100), which represents the initial position of the axes of the specimen. S_1

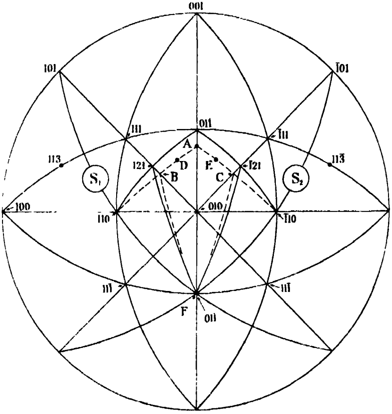


FIG. 11.

and S_2 are the two possible slip planes whose poles are shown at $(\bar{1}11)$ and (111) respectively. The directions of slip are the normals (110) on S_1 and $(\bar{1}10)$ on S_2.

As slipping proceeds the direction of the axis of the specimen moves relatively to the crystal axes in the direction AB in one part of the specimen and AC in the other. At any given extension, say, 30 per cent., the two points representing the axis of the specimen in the two parts will move to D and E respectively where AD equals AE.

As slipping proceeds the points representing the axis of the specimen move along great circles towards the poles of the planes (110) and $(\bar{1}10)$ till they reach the points B and C, where these great circles cut the dodecahedral planes (101) and $(\bar{1}01)$. At this stage double slipping begins in each part into which the specimen is divided and the points representing the axis of the specimen move very slowly towards the poles of the planes (121) and $(\bar{1}21)$.

The amount of extension which the specimen sustains before the double slipping begins depends on the position of the point A. Its maximum value is 73 per cent. when A coincides with the pole (011). Referring to Dr. Elam's diagram figs. 7 and 8, it will be seen that at the stage at which her X-ray analysis was carried out, the axis of the specimen was very close to the normal to the plane (121) in one part and $(\bar{1}21)$ in the other. It is improbable that the total extension of the specimen was much greater than 70 per cent. (though it was not measured). Consequently, it seems likely that A was close to the pole (011) as is shown in the diagram, fig. 10.

At the beginning of the test the plane of union of the two layers of the specimen must in any case have been parallel to the cubic plane (100). The line of particles represented by F (fig. 11), which lie along the normal to the plane $(01\bar{1})$, is in the plane of junction and also in both slip planes S_1 and S_2, so that this plane of particles remains normal to the plane $(01\bar{1})$ in both parts of the specimen as the specimen extends. The plane of union passes through this normal and through the axis of the specimen. At the moment when double slipping begins therefore, the plane of union is represented by the dotted circle BF in one part and the circle CF in the other.

Since the point A is close to the pole (011) and the point B therefore close to the pole (121), the plane of junction nearly coincided with the plane passing through the poles (121) and $(01\bar{1})$ at the time the double slipping began in both parts of the specimen. This plane, which is represented by indices $(11\bar{3})$, is marked in fig. 11. The corresponding plane for the other part of the specimen is that passing through the poles $(\bar{1}21)$ and $(01\bar{1})$. These are in fact the orientations which Dr. Elam finds by X-ray analysis.

It appears, therefore, that if the laws previously discovered regarding the distortion of aluminium crystals are applied to the case of a tensile specimen whose longitudinal axis lies in a cubic crystal plane and close to the normal to an octahedral plane, "mechanical twins" of the type discovered by Dr. Elam might have been predicted as possible, and the orientation of the two sets of crystal axes would bear the same relationship to the plane of union and the longitudinal axis of the specimen as that actually found.

The Change in Lattice Spacing at a Crystal Boundary.

By J. E. Lennard-Jones and Beryl M. Dent, The University, Bristol.

(Communicated by S. Chapman, F.R.S.—Received April 27, 1928.)

§ 1. *Introduction.*

Some recent experiments of Davisson and Germer* on the scattering of electrons by a crystal have drawn attention to the conditions at a crystal boundary. In interpreting their results in terms of de Broglie waves, the authors have to postulate a contraction of the crystal lattice at the surface—in some cases of as much as 30 per cent.† It is, therefore, important that other independent methods should be devised to indicate what change (if any) takes place at a crystal surface. Unfortunately X-rays are unable to help in this respect as several hundred atomic layers are necessary to produce a Bragg reflection pattern; nor is theory able to provide an answer in the case of metallic crystals such as are used by Davisson and Germer because little is known of the forces which hold a metal together. Certain other cases can, however, be considered theoretically, and these may indicate the order of the effect to be expected in metallic crystals.

In this paper we consider the change in spacing at a (100) boundary of a crystal of the NaCl type, this boundary being considered because it is a natural plane of crystal. Two effects are to be anticipated: (1) a change of spacing between planes at the boundary; (ii) a change of spacing between atoms in the surface layer itself. These effects are considered separately and independently. Actually, of course, each affects the other, but this is assumed to be a correction of the second order.

---

* Davisson and Germer, 'Nature,' vol. 119, p. 558 (1927); 'Phys. Rev.,' vol. 30, p. 705 (1927). The experiments of G. P. Thomson ('Roy. Soc. Proc.,' A, vol. 117, p. 600 (1928)) are not affected in the same way by surface conditions owing to the high speed of the electrons.

† See note at end of paper.

§ 2. *The Attractive Force between the Top Layer and the Rest of the Crystal.*

In calculating the change in spacing between the first and second layers, we *neglect the change between lower layers. The subsequent calculations will indicate how far this assumption is justified. We also suppose the spacing in the layers themselves to be the same as that in the interior of the crystal.*

The electrostatic force between an ion in the surface layer and the rest of the crystal can be calculated by a method due originally to Madelung,[*] as has been done in detail for crystals of this type in a recent paper.[†] Under the influence of this field (E^c) the surface ions are polarised, and if we suppose, as usual, that polarisation of each ion can be represented by a dipole whose strength is proportional to the exciting field, then the array of dipoles at the surface itself produces another field E^d. The exciting field at an ion is then the sum of E^c and E^d, so that the dipole strengths are given by

$$p_r = \alpha_r (E_r^c + E_r^d), \qquad (r = 1, 2) \tag{2.01}$$

where α_1 and α_2 are constants characteristic of the two ions present.[‡] In the case considered both E^c and E^d are normal to the crystal boundary. For definiteness, we suppose E^c to be positive when a positive charge is repelled from the surface, and a dipole is regarded as positive when its positive pole is uppermost.

The potential energy of each dipole is easily seen to be

$$\Phi_r = -\tfrac{1}{2}\alpha_r (E_r^c + E_r^d)^2, \qquad (r = 1, 2) \tag{2.02}$$

so that the force on each away from the surface is

$$-\frac{\partial \Phi_r}{\partial z} = \alpha_r (E_r^c + E_r^d)\left(\frac{\partial E_r^c}{\partial z} + \frac{\partial E_r^d}{\partial z}\right), \tag{2.03}$$

the axis of z being normal to the surface. The total force on two ions in the surface, one of each kind, is therefore

$$F(z) = \Sigma \left\{ v_r e (E^c + E^d)_r + \alpha_r (E^c + E^d)_r \frac{\partial}{\partial z}(E^c + E^d)_r \right\}, \tag{2.04}$$

where v_r is the valency of each ion.

---

[*] Madelung, 'Phys. Z.,' vol. 19, p. 524 (1918); Born, "Atomtheorie des festen Zustandes," 2nd edn., § 37 (1923).

[†] Lennard-Jones and Dent, 'Trans. Faraday Soc.,' February, 1928.

[‡] *Cf.* Born and Heisenberg, 'Z. f. Physik,' vol. 23, p. 388 (1924).

Now the field E^d, being due to the array of dipoles, must be expressible in terms of the dipole moments p_1 and p_2, and therefore in terms of E^c. Referred to rectangular axes in the surface, the co-ordinates of the dipoles are $(\frac{1}{2}la, \frac{1}{2}ma)$, where l and m are integers and a is the distance between two ions of the same kind on a cubic axis. One set of dipoles is at the points for which $l+m$ is odd, and the other at the points for which $l+m$ is even. For definiteness we will suppose that the p_1 dipoles are at the odd points and the p_2 dipoles at the even points.

The field at the origin due to the p_1 dipoles is directed towards the crystal and is given by $-(2/a)^3 S_1 p_1$, where

$$S_1 = \Sigma(l^2+m^2)^{-3/2}, \qquad (l+m = \text{odd})$$

—a summation over all values of l and m, positive and negative, for which $l+m$ is odd. The field at the same point due to the p_2 system of dipoles is given by $-(2/a)^3 S_2' p_2$, where S_2' is a similar summation over values of l and m for which $l+m$ is even, omitting the term $l=m=0$.

We have then

$$E_2^d = -(2/a)^3(p_1 S_1 + p_2 S_2'), \tag{2.05}$$

and similarly

$$E_1^d = -(2/a)^3(p_1 S_2' + p_2 S_1). \tag{2.06}$$

Hence

$$\left.\begin{aligned} p_1 &= \alpha_1 E_1 = \alpha_1\{E_1^c - (2/a)^3(p_1 S_2' + p_2 S_1)\} \\ p_2 &= \alpha_2 E_2 = \alpha_2\{E_2^c - (2/a)^3(p_1 S_1 + p_2 S_2')\} \end{aligned}\right\}. \tag{2.07}$$

These two equations determine p_1 and p_2 in terms of the field E^c due to the crystal; in fact, we may write

$$p_1 = \alpha_1 E_1^c \eta_{11} + \alpha_2 E_2^c \eta_{12}, \quad p_2 = \alpha_1 E_1^c \eta_{21} + \alpha_2 E_2^c \eta_{22}, \tag{2.08}$$

where the η's involve only known (or calculable) constants. It follows that

$$E_1^c + E_1^d = E_1^c \eta_{11} + (\alpha_2/\alpha_1) E_2^c \eta_{12}, \tag{2.09}$$

$$E_2^c + E_2^d = (\alpha_1/\alpha_2) E_1^c \eta_{21} + E_2^c \eta_{22}, \tag{2.10}$$

so that the total force on a cell in the surface away from the crystal is

$$\begin{aligned} F(z) = {} & v_1 e\{E_1^c \eta_{11} + (\alpha_2/\alpha_1) E_2^c \eta_{12}\} + v_2 e\{(\alpha_1/\alpha_2) E_1^c \eta_{21} + E_2^c \eta_{22}\} \\ & + \{\alpha_1 E_1^c \eta_{11} + \alpha_2 E_2^c \eta_{12}\}\left\{\eta_{11}\frac{\partial E_1^c}{\partial z} + (\alpha_2/\alpha_1)\eta_{12}\frac{\partial E_2^c}{\partial z}\right\} \\ & + \{\alpha_1 E_1^c \eta_{21} + \alpha_2 E_2^c \eta_{22}\}\left\{(\alpha_1/\alpha_2)\eta_{21}\frac{\partial E_1^c}{\partial z} + \eta_{22}\frac{\partial E_2^c}{\partial z}\right\}. \end{aligned} \tag{2.11}$$

For a crystal of the NaCl type, where $v_1 = 1$, $v_2 = -1$, $E_2{}^c = -E_1{}^c$, we find

$$F(z) = e\{\eta_1 + \eta_2\} E_1{}^c + \{\alpha_1\eta_1{}^2 + \alpha_2\eta_2{}^2\} E_1{}^c\, \partial E_1{}^c/\partial z, \tag{2.12}$$

where

$$\eta_1 = \eta_{11} - (\alpha_2/\alpha_1)\,\eta_{12} = \{1 + \alpha_2 (2/a)^3 (S_1 + S_2')\}/\Delta, \tag{2.13}$$

$$\eta_2 = -(\alpha_1/\alpha_2)\,\eta_{21} + \eta_{22} = \{1 + \alpha_1 (2/a)^3 (S_1 + S_2')\}/\Delta, \tag{2.14}$$

and

$$\Delta = 1 + (\alpha_1 + \alpha_2)(2/a)^3 S_2' + \alpha_1\alpha_2 (2/a)^6 (S_2'^2 - S_1{}^2). \tag{2.15}$$

We note that

$$p_1 = \alpha_1 E_1{}^c \eta_1, \qquad p_2 = \alpha_2 E_2{}^c \eta_2, \tag{2.16}$$

so that η_1 and η_2 give the fractional increases in the field at the ions due to the presence of the other dipoles.

The force $F(z)$ on a unit cell in the surface is now expressed in terms of $eE_1{}^c$, which is the force on a positive ion in the surface due to all the charges in the crystal except those in the surface layer. This has been worked out in a recent paper.* We find

$$E_1{}^c = -\frac{2\pi e}{a^2} \sum_{l,m}' \frac{\exp(-2\pi(l^2+m^2)^{\frac{1}{2}} z/a)}{1 + \exp(-\pi(l^2+m^2)^{\frac{1}{2}})} \{1 + (-1)^{l+m} - (-1)^l - (-1)^m\} \tag{2.17}$$

which may be written

$$E_1{}^c = -(8\pi e/a^2) \sum_{l,m}' c_{lm} f_{lm}(z/a) = -(32\pi e/a^2)\, s_1(z/a).$$

A number of values of s_1 have been calculated in the paper just referred to, and others required in this paper are given in the table below. We require as well the value of $\partial E_1{}^c/\partial z$ and for this we find

$$\frac{\partial E_1{}^c}{\partial z} = \frac{16\pi^2 e}{a^3} \sum_{l,m}' (l^2 + m^2)^{\frac{1}{2}} c_{lm} f_{lm}(z/a) = \frac{64\pi^2 e}{a^3} s_2(z/a).$$

Table I.—Values of Certain Summations.

z/a	0 42	0·45	0·46	0·48	0·49	0·50
$s_1 . 10^2$	2·416	1·840	1·681	1·403	1·282	1·172
$s_2 . 10^2$	3 505	2·650	2·416	2·011	1·835	1·675

There remain only the summations S_1 and S_2' to be evaluated. Following a method due to one of the present authors and A. E. Ingham,† Topping‡ has shown that $S = S_1 + S_2' = 9{\cdot}0336$. From this result the individual values of S_1 and S_2' can easily be deduced. The lattice points included in S_2' form a square array of side $\sqrt{2}$, so that $S_2' = (2)^{-3/2} S$ and hence

---

* Lennard-Jones and Dent, 'Trans. Faraday Soc.,' *loc. cit.*

† 'Roy. Soc. Proc.,' A, vol. 107, p. 636 (1925).

‡ Topping, 'Roy. Soc. Proc.,' A, vol. 114, p. 67 (1927).

$$S_1 = \{1 - (2)^{-3/2}\}(9\cdot0336)\,; \quad S_2' = (2)^{-3/2} \quad (9\cdot0336). \tag{2.18}$$

The force F_s may now be regarded as a known function of z, since the constants α_1 and α_2 are atomic constants, which we may suppose known.

§ 3. *Intrinsic Repulsive Fields.*

We suppose the ions to repel each other according to an inverse power of the distance. The forces between like and unlike ions may then be written $\lambda_{11}r^{-n_{11}}$, $\lambda_{22}r^{-n_{22}}$ and $\lambda_{12}r^{-n_{12}}$. If we suppose, as above, that the distance between the layer and the second layer is z, and the distance between other layers is $a/2$, then the normal force on a cell in the surface containing two ions is

$$H(z) = (2/a)^{n_{11}}\lambda_{11}h_1^{(n_{11})}(z) + 2\,(2/a)^{n_{12}}\lambda_{12}h_2^{(n_{12})}(z) + (2/a)^{n_{22}}\lambda_{22}h_1^{(n_{22})}(z), \tag{3.01}$$

where

$$h_1^{(s)}(z) = \Sigma_1 \{2\,(z/a) + n\}\,\{l^2 + m^2 + (2\,(z/a) + n)^2\}^{-(s+1)/2}, \tag{3.02}$$

$$h_2^{(s)}(z) = \Sigma_2 \{2\,(z/a) + n\}\,\{l^3 + m^2 + (2\,(z/a) + n)^2\}^{-(s+1)/2}, \tag{3.03}$$

the first summation being over odd values of $(l + m + n)$ and the second over even values.

A method of evaluating summations such as these has been given in a recent paper.* It consists in calculating by direct summation the contribution of the ions within a certain spherical cap and in replacing the rest by a continuum, each volume element of which repels according to the appropriate law. Some of the values needed below are given in Table II.

Table II.—Values of $h_1^{(s)}(z)$ and $h_2^{(s)}(z)$.

z/a.	0·42.	0·45.	0·46.	0·48.	0·49.	0·50.
$h_1^{(9)}(z)$	0·2405	0·1916	0·1774	0·1518	0 1403	0 1295
$h_1^{(10)}(z)$	0 1819	0·1406	0 1288	1·1080	0·0988	0 0904
$h_1^{(11)}(z)$	0 1383	0·1037	0·0941	0·0773	0·0700	0·0634
$h_2^{(9)}(z)$	4·833	2 607	2·143	1·467	1·221	1·021
$h_2^{(10)}(z)$	5 734	2·883	2·316	1·517	1·236	1·011
$h_2^{(11)}(z)$	6·816	3·195	2·510	1·574	1·256	1 006

---

* Lennard-Jones and Dent, 'Trans. Faraday Soc.,' *loc. cit.*

§ 4. *Change of Crystal Spacing at the Surface.*

The equation to determine z, the distance between the top layer and the second layer, is

$$F(z) + H(z) = 0, \tag{4.01}$$

where $F(z)$ is given by equations (2.12), (2.13), (2.14), (2.15) and (2.17). The force constants involved in $H(z)$ have been determined in a series of recent papers.* Those referring to certain sodium and potassium crystals are reproduced in Table III, the appropriate values of n being given in brackets.

Table III.—The Force Constants and Force Indices of certain Ions†
(λr^{-n}).

Na—F	1·04 (11)	K—F	1·45 (10)	F—F	3·81 (11)
Na—Cl	2·12 (10)	K—Cl	1·59 (9)	Cl—Cl	4·49 (9)
Na—Br	2·75 (10)	K—Br	1 95 (9)	Br—Br	30·7 (10)
Na—I	14·1 (11)	K—I	11·5 (10)	I—I	256·5 (11)
Na—Na	0·232 (11)	K—K	0·481 (9)		

The utility of these force constants can be illustrated by calculating the interatomic distances in the interior of crystals, when the results can be compared with those determined experimentally. It can be shown that the equation to determine $a/2$ in a crystal of the NaCl type is‡

$$\lambda_{11}A''_{n_{11}-1}x^{n_{11}} + \lambda_{22}A''_{n_{22}-1}x^{n_{22}} + 2\lambda_{12}A'_{n_{12}-1}x^{n_{12}} = e^2v^2sx^2, \tag{4.02}$$

where A′, A″ and s are certain summations over the lattice points of the crystal,§ and v, as before, is the valency of the ions. The value of s for crystals of the NaCl type is 3·495.|| In this calculation the polarisation of the ions need not be considered owing to the symmetrical arrangement of the ions. The results of the calculations, together with the observed values are given in Table IV.

---

* 'Roy. Soc. Proc.,' A, vol. 106, pp. 441 and 463 (1924), vol. 107, pp. 157, 636 (1925); vol. 109, pp. 476 and 584 (1925); vol. 112, pp. 214 and 230 (1926); vol. 115, p. 334 (1927).

† (The force constants are in such units as give the force in dynes when the unit of length is the Ångström).

‡ Lennard-Jones and Taylor, 'Roy. Soc. Proc.,' A, vol. 109, p. 476 (1925).

§ Lennard-Jones and Ingham, 'Roy. Soc. Proc.,' A, vol. 107, p. 636 (1925).

|| Madelung, *loc. cit.*

Table IV.—Calculated and Observed Interatomic Distances in Crystals.*

	F.		Cl.		Br.		I.	
	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.
Na	2·30	2·31	2·86	2·81	2·99	2·97	3·19	3·23
K	2·65	2·66	3·13	3 14	3·24	3 28	3·47	3 52

Before considering the general equation (4.01) we consider the simplified form when polarisation at the surface is neglected. The equation then becomes

$$2eE_1^c(z) + H(z) = 0, \qquad (4.03)$$

where $E_1^c(z)$ is given by equation (2.17). The solution of this equation for the series of crystals just considered is given in Table VI. The results have been compared with the theoretically calculated values for the interior given in the table above, which is the only logical course to take. Both sets of results are then a direct consequence of the same force constants.

In solving the complete equation (4.01), the following values of the polarisation constants have been used (Table V).† They cannot be regarded as known with accuracy but they suffice to indicate the order of magnitude of the contraction at the surface.

Table V.—Polarisation Constants of certain Ions.

	Na.	K.	F.	Cl.	Br.	I.
$\alpha\ 10^{24}$	0·21	0·87	0·99	3·05	4·17	6 28

The values of the quantities η_1 and η_2 given in equations (2.13) and (2.14) were then found. Usually $\eta_1 > 1$ and $\eta_2 < 1$, but their sum is approximately constant (2·25 to 2·41). The former gives the fractional change in the field at the positive ions due to the presence of the other dipoles, the latter the change at the negative ions. The value of $F(z)$ is found to be about 20 to 30 per cent. greater than $2eE_1^c(z)$ (the value of $F(z)$ when polarisation is

---

* The observed values are from W. H. and W. L. Bragg, 'X-rays and Crystal Structure,' p. 306 (1924).

† Born and Heisenberg, 'Z. f. Physik,' vol. 23, p. 388 (1924).

neglected) when $z = a/2$, and about 30 to 40 per cent. greater when $z = (0 \cdot 42)a$. This illustrates the possible importance of polarisation at the surface.

Equation (4·01) was solved by graphical methods with the results given in Table VI. The average percentage decrease is seen to be of the order of 5 per cent. Since the crystals are ionic and the attractive forces are purely electrostatic, it is unlikely that the contraction will be greater than this in crystals of non-polar type.

Table VI.—The Percentage Decrease in the Interplanar Spacing at a Crystal Surface.

Percentage decrease.	NaF.	NaCl.	NaBr.	NaI.	KF.	KCl.	KBr.	KI.
Neglecting polarisation	0·6	0·1	0·8	0·8	0·7	0·2	0·5	0·0
Including polarisation	3·4	5·0	6·2	5·6	4·8	6·9	8·0	6·2

Seeing that the contraction is due almost entirely to the polarisation of the surface layer, it is clear that the contraction of the spacing between the second and third layers will be much smaller, because the second layer will be polarised not only by the layers below it but also by the layer above it. The two effects will be in opposite senses and will almost completely compensate each other. We conclude, therefore, that the contraction of the lattice between layers other than the first and second is negligible.

§ 5. *The Effect of Van der Waals' Attractive Fields.*

In a recent paper* it has been shown by the present authors that the effect of Van der Waals' attractive forces is important outside a crystal surface and it is therefore necessary to investigate the effect of these forces on the surface layer. Unfortunately, insufficient is known of the forces of this type between ions to make an exact calculation possible, though some information has been found about the magnitude of such forces between *inert gas atoms*,† which can be used to estimate the order of magnitude of the effect. In so doing we must remember that a similar correction must be introduced into the calculations of the spacing in the interior. The Van der Waals' attractive forces are due to the electronic systems of the ions apart altogether from the net electrostatic charges of the ions. Their magnitude will (presumably) be of the same order as that of the inert gas atoms of similar electronic structure. Thus, since the Van der Waals' forces between neon atoms are known, we may suppose that all the ions in NaF behave in this respect as though they were neon atoms. Representing this force by the law μr^{-m}, we find for neon* $m = 5$, $\mu = 1\cdot 72, 10^{-44}$. The appropriate equation to determine the internal spacing constant is

---

* Lennard-Jones and Dent, 'Trans. Faraday Soc.,' *loc. cit.*

† Lennard-Jones and Cook, 'Roy. Soc. Proc.,' A, vol. 112, p. 214 (1926).

$$\lambda_{11}\, A''_{n_{11}-1}\,(x)^{n_{11}} + 2\lambda_{12}\, A'_{n_{12}-1}\,(x)^{n_{12}} + \lambda_{22}\, A''_{n_{22}-1}\,(x)^{n_{22}} = (3\cdot 495)\, e^2\,(x)^2 + 2\mu A_{m-1}\,(x)^m,$$

though the equation simplifies in the case of NaF since $n_{11} = n_{12} = n_{22} = 11$.

The value of $a/2$ for NaF determined by this equation proves to be 2·28 Å instead of 2·30 Å. when μ is neglected.

The equation (4.01) to determine the spacing at the boundary must similarly be modified by the addition of a term

$$-\,2\mu\,(2/a)^m\,\{h_1^{(m)}(z) + h_2^{(m)}(z)\}$$

to the left hand side. For $m = 5$. we find

z/a	0·42.	0 45.	0·46.	0·48.	0 49.	0·50.
$h_1^{(5)}(z) + h_2^{(5)}(z)$	3·749	2·943	2·733	2·374	2 221	2 083

using the methods referred to above (§ 3) and described in an earlier paper.†

Using the calculated value $a/2 = 2\cdot 28$ Å as the spacing in the interior, we find the following percentage decrease at the boundary :—

$$\text{NaF}\left\{\begin{array}{lll}\text{Neglecting polarisation} & .. & 0 \text{ per cent.}\\ \text{Including polarisation} & .. & 3 \quad ,,\end{array}\right.$$

The effect of including Van der Waals' forces throughout the calculations has therefore scarcely modified the results.

A similar investigation for KCl, assuming the Van der Waals' forces to be the same as those of argon (to which all the ions are similar), shows the following percentage decrease at the boundary :—

$$\text{KCl}\left\{\begin{array}{lll}\text{Neglecting polarisation} & .. & 0 \text{ per cent.}\\ \text{Including polarisation} & .. & 6 \quad ,,\end{array}\right.$$

so that the results given in the previous paragraph are scarcely affected.

---

* Lennard-Jones and Cook, 'Roy. Soc. Proc.,' A, vol. 112, p. 214 (1926).

† Lennard-Jones and Dent, 'Trans. Faraday Soc.,' *loc. cit.*

§ 6. *Contraction in the Surface Layer.*

In addition to the change in interplanar spacing at the boundary, we may expect a change in the spacing in the top layer itself. A full investigation of this effect, taking into account the influence of the rest of the crystal, is not attempted here, but some indication of its magnitude may be obtained by considering the spacing in a single isolated layer. This will give an upper limit to the effect.

Consider a plane of positive and negative ions arranged alternately in a square array as in the (100) plane of a NaCl crystal. Let $a/2$ be the distance between neighbouring ions. A square cell of side $a/\sqrt{2}$ obtained by joining four neighbouring ions of the same sign may be regarded as containing two ions, one of each sign. Let the potential energy of the cell due to the electrostatic forces be $\phi^{(e)}$ and that due to the intrinsic repulsive forces be $\phi^{(n)}$, so that the total potential energy is

$$\phi = \phi^{(e)} + \phi^{(n)}.$$

In this case we do not need to consider polarisation as each ion is surrounded symmetrically by ions of opposite charge.

The potential energy of an ion in a row of equidistant ions, alternately negative and positive, due to all the rest is

$$\phi_1^{(e)} = e \sum_{l=0}^{\infty} \left\{\frac{2(-e)}{(l+\frac{1}{2})a} + \frac{2e}{(l+1)a}\right\} = -\frac{4e^2}{a}\log_e 2.$$

The potential energy of a similar row at a point outside it is given by*

$$\psi(x, r) = \frac{2}{a} \sum_{-\infty}^{+\infty}{}'_n \left\{-eK_0\left(\frac{2\pi nr}{a}\right) e^{\frac{2\pi inx}{a}} + eK_0\left(\frac{2\pi nr}{a}\right) e^{\frac{2\pi in(x-a/2)}{a}}\right\}.$$

where K_0 is a Hankel function,† (x, r) are the cylindrical co-ordinates of the point considered referred to an origin at one of the negative charges in the row and the axis of x is taken along the row of charges. Hence the potential energy of a positive ion due to all the other ions of a plane square array except those in the same row is

$$\phi_2^{(e)} = \frac{4e^2}{a} \sum_{-\infty}^{+\infty}{}'_n \Sigma_s (-1)^s \{K_0(\pi sn) - (-1)^n K_0(\pi sn)\}.$$

Using the fact that $K_0(\pi n) = K_0(-\pi n)$ we have

$$\phi_2^{(e)} = \frac{16e^2}{a} \sum_{1}^{\infty}{}_n \sum_{1}^{\infty}{}_s (-1)^s K_0(\pi s(2n+1)).$$

---

* Born, "Atomtheorie des festen Zustandes," 2nd edn., p. 718 (1923).

† Jahnke and Emde, 'Funktionen tafeln,' table XIV, p. 135 (1909). Topping and Ludlam, 'Phil. Trans.,' A, vol. 226, p. 135 (1927).

This is a rapidly converging series (thus $K_0(\pi) = 2\cdot95 \,.\, 10^{-2}$, $K_0(2\pi) = 0\cdot09 \,.\, 10^{-2}$, $K_0(3\pi) = 0\cdot003 \,.\, 10^{-2}$) and is easily summed. We thus have

$$\phi^{(e)} = 2\left(\phi_1^{(e)} + \phi_2^{(e)}\right) = -\frac{3\cdot232e^2}{(a/2)}.$$

The potential energy of a cell due to the intrinsic repulsive forces is easily seen to be

$$\phi^{(n)} = \frac{\lambda_{11}}{(n_{11}-1)}\left(\frac{2}{a}\right)^{n_{11}-1} S''_{n_{11}-1} + \frac{2\lambda_{12}}{(n_{12}-1)}\left(\frac{2}{a}\right)^{n_{12}-1} S'_{n_{12}-1} + \frac{\lambda_{22}}{(n_{22}-1)}\left(\frac{2}{a}\right)^{n_{22}-1} S''_{n_{22}-1},$$

where

$$S'_n = \sum_{(l+m)\text{ odd}} (l^2+m^2)^{-n/2},$$

$$S''_n = \sum_{(l+m)\text{ even}}{}' (l^2+m^2)^{-n/2}.$$

The values of these summations for $n = 8$, 9 and 10 are given in Table VII.

Table VII.—Values of the Summations S'_n and S''_n.

n.	8	9.	10.
S'_n	4 016	4 008	4 004
S''_n	0 2676	0 1852	0 1292

In equilibrium $\partial\phi/\partial a = 0$ and so the equation to determine $a/2$ is

$$\lambda_{11} S''_{n_{11}-1}\, x^{2-n_{11}} + 2\lambda_{12} S'_{n_{12}-1}\, x^{2-n_{12}} + \lambda_{22} S''_{n_{22}-1}\, x^{2-n_{22}} = (3\cdot232)\, e^2$$

The interatomic distances of isolated (100) planes of the crystals considered earlier in this paper, determined by this equation, are given in Table VIII. These are compared with the values calculated for the interior of the crystal.

Table VIII.—Interatomic Distances in Isolated (100) Planes.

Crystal spacing.	NaF.	NaCl.	NaBr.	NaI.	KF.	KCl.	KBr.	KI.
In isolated (100) plane	2 20	2·69	2·80	3 00	2·52	2·94	3·04	3·30
In interior of crystal (calc.)	2 30	2·86	2·99	3 19	2 65	3·13	3·24	3·47
Percentage decrease	4½	6	6½	6	5	6	6	5

The decrease is seen to be of the order of 5 or 6 per cent.* The actual decrease in the surface layer of a crystal will, of course, be less than this owing to the restraining influence of the lower layers. The general tendency will, however, be towards contraction and this may lead to cracking at the surface.

§ 7. *Surface Tension.*

Since there is a tendency for the surface layer to contract, it follows that a lateral force or tension is necessary to maintain the same spacing at the boundary as in the interior.

If T is the tension per unit length, we have

$$T = \frac{\partial \phi}{\partial A} = \frac{1}{2\rho}\frac{\partial \phi}{\partial \rho},$$

where A is the area and ρ the side of a unit cell. Hence

$$T = \frac{1}{2\rho}\frac{\partial}{\partial r}(\phi^{(e)} + \phi^{(n)})\frac{\partial r}{\partial \rho} = \frac{1}{4r}\frac{\partial}{\partial r}(\phi^{(e)} + \phi^{(n)}),$$

where r is the shortest distance between ions and $\rho = \sqrt{2}\, r$. We thus have

$$T = \tfrac{1}{4}\left\{\frac{(3\cdot 232)\, e^2}{r^3} - \frac{\lambda_{11} S''_{n_{11}-1}}{r^{n_{11}+1}} - \frac{2\lambda_{12} S'_{n_{12}-1}}{r^{n_{12}+1}} - \frac{\lambda_{22} S''_{n_{22}-1}}{r^{n_{22}+1}}\right\},$$

The variation of T with r for a few cases is shown in the figure. In each case T increases linearly at first as is to be expected; in fact the curves show the range for which Hooke's law in this particular case is valid. After reaching a maximum, T decreases asymptotically as $1/r^3$, this being due to the electrostatic forces of attraction.

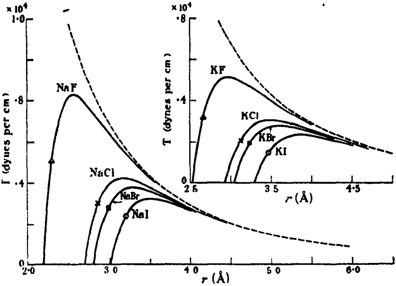
The maximum value for NaF occurs when $r = 2\cdot 57$ A and then $T_{max} =$ 8200 dynes per centimetre. The values of T corresponding to $r = a/2$ are marked in the figure. Approximate values are given in the following table.

Table IX.—Surface Tension of Certain Crystals (dynes per centimetre).

	NaF.	NaCl.	NaBr.	NaI.	KF.	KCl.	KBr.	KI.
T	5100	3000	2800	2350	3200	2050	1950	1500

---

* Zwicky has, in the course of a calculation of the breaking strength of NaCl, given an estimate of the interatomic distance in a single plane of Na and Cl ions. Assuming all the ions to repel according to the same law, he finds the decrease to be 5·37 per cent. ('Phys. Z.,' vol. 24, p. 131 (1923)).



The Relation between Tension and Extension in the (100) planes of certain crystals.

§ 8. *Summary.*

(1) The contraction of the lattice at the (100) boundary of crystals of the NaCl type is shown to be confined almost entirely to the top layer and to be of the order of 5 per cent.

(2) An upper limit is found for the decrease in the interatomic spacing in the surface layer. This also is of the order of 5 per cent.

(3) The surface tension of a number of crystals of the NaCl type is calculated.

[*Note added 2nd October*, 1928.—It has now become clear that the results of Davisson and Germer (*loc. cit.*) can be explained by supposing the electrons to be refracted in the crystal in an optical sense with an index of refraction greater than unity*—a more acceptable hypothesis than that of excessive contraction at the surface. This paper provides independent evidence against the contraction theory because it shows that the contraction is too small and localised to produce an observable effect on the reflection pattern of an electron stream.]

---

* Davisson and Germer, 'Proc Nat. Acad Sci.,' vol 14, pp. 317 and 619 (1928); Bethe, 'Naturwiss,' vol. 15, p. 787 (1927), vol. 16, p. 333 (1928).

The Influence of the Earth's Magnetic Field on Electric Transmission in the Upper Atmosphere.

By S. Goldstein, B.A., St. John's College, Isaac Newton Student in the University of Cambridge.

(Communicated by Sir Joseph Larmor, F.R.S. Received June 1, 1928.)

The Ionic Theory of Transmission in the Upper Atmosphere.

In the 'Philosophical Magazine' for December, 1924, Sir Joseph Larmor showed how wireless waves can be transmitted to great distances, round the protuberance of the curved earth, and without excessive damping, if the transmission takes place in an ionised region high in the ultra-rarefied upper atmosphere, in which the number of effective ions increases upwards.

Under the influence of the waves the ions oscillate, and thus produce a current which must be added, in the electrodynamic equations of the exciting wave, to the æthereal displacement current. The velocity of propagation is thus altered to c', where

$$c'^{-2} = c^{-2}\,(1 - 4\pi \mathrm{N} e^2 c^2/mp^2),$$

c being the velocity of light, N the number of effective ions per unit volume, $p/2\pi$ the frequency of the electric oscillations, e the ionic charge and m the ionic mass. (All units are electrodynamic.) If N increases upwards, c' increases upwards, and the rays will be bent downwards, with a curvature $c'^{-1}\,\partial c'/\partial h$. In this way we see that it may be possible for a ray to bend round and follow the curvature of the earth.

The energy lost by secondary radiation from the ions is negligible, since the oscillations are slow compared with those of light. Dissipation is due mainly to collisions of the ions with the molecules of the atmosphere; and it is in the ultra-rarefied upper atmosphere, and there only, that the dissipation due to this cause will not be excessive when the gradient of N is sufficient to give the correct curvature. For there the mean free path is long; the amplitude of the oscillations fits easily into this mean free path; and the mean time between successive collisions may be a fairly large multiple of the period of an oscillation.

The reason why the comparatively small number of effective ions in the upper atmosphere produces an appreciable effect on wireless waves, but practically no effect on the short waves of light, is that the effect is inversely proportional to the square of the frequency, as shown by the formula above.

Thus it appears that in long distance wireless communication a sheaf of rays is concentrated in a layer in an ionised region high in the ultra-rarefied upper atmosphere, the layer being that in which the gradient of N will just bend the rays round the earth. Local irregularities in the vertical distribution of N are to be expected; the rays will not follow the curvature of the earth exactly, but may sometimes be moving slightly upwards and sometimes slightly downwards; and if the changes in the distribution of N are abrupt, as they may be expected to be across a sunset or sunrise band, the rays may be bent upwards into space or downwards to the earth so far that they cannot recover the horizontal direction.

Finally, two other questions that present themselves were answered in the affirmative by Larmor. Can there be concentrated into the layer of effective transmission sufficient energy to be showered down to receivers all over the earth? and is it possible for a thin sheaf of rays to travel concentrated in a layer without spreading sideways?

The presence of an ionised region high in the ultra-rarefied upper atmosphere, and its effect on wireless communication, have received confirmation from the results of several experimenters, notably Appleton and his co-workers,* who have shown that over short distances transmission takes place not only by means of a ground ray, but also by means of an atmospheric ray that is gradually reflected downwards in the ionised region. For a full description of experimental methods and results, reference should be made to the papers cited below.

The Influence of the Earth's Magnetic Field.

Prof. E. V. Appleton has pointed out† that when the ions oscillate they will come under the influence of the earth's magnetic field, and that this will have an appreciable effect upon the circumstances of propagation. In class lectures at Cambridge Sir Joseph Larmor has considered the effect of the earth's magnetic field on the propagation of wireless waves in two special cases—when the imposed magnetic field is along the direction of propagation of the wave, and the wave is circularly polarised; and when the imposed magnetic field is perpendicular to the direction of propagation and the wave is plane polarised with the magnetic vector along the direction of the imposed magnetic field. He pointed out that for any given value of N and of the magnetic field there are certain values of p (*i.e.*, certain wave-lengths) for which $(c/c')^2$ would be negative, so that such waves will not travel; and that this implies that the rays in these cases will never penetrate to the region in which N has the assumed value, but will be bent round and come down again to earth; and this no matter how near the vertical such a ray enters the ionised region. On the other hand, the knowledge that waves of a certain wave-length are transmitted to great distances will enable us to set limits to the possible value of N in the layer of effective transmission.

---

* 'Roy. Soc. Proc.,' A, vol. 109, p. 621 (1925); vol. 113, p. 450 (1926); vol. 115, p. 291 (1927), and vol. 117, p. 576 (1928). See also Smith-Rose and Barfield, 'Roy. Soc. Proc.,' A, vol. 107, p. 587 (1925), vol. 110, p. 580 (1926); vol. 116, p. 682 (1927); Round, Eckersley, Tremellen and Lunnon, 'J. Inst. Elec. Eng.,' vol. 63, p. 933 (1925); Hollingworth, 'J. Inst. Elec. Eng.,' vol. 64, p. 579 (1926); Breit and Tuve, 'Phys. Rev.,' vol. 28, p. 554 (1926).

† 'Proc. Phys. Soc.,' vol. 37 (1925). See also Nichols and Schelling, 'Bell System Tech. J.,' vol. 4, p. 215 (1925); Breit, 'Proc. Inst. R. Eng.,' vol. 15, p. 709 (1927).

Below, formulæ are developed for the general problem, when the magnetic field is oblique to the direction of propagation. The formulæ for the two special cases are in agreement.

A Wave-train in an Ionised Medium with Imposed Magnetic Field Oblique to the Direction of Propagation.

Consider a plane wave-train travelling along the negative direction of the axis of z.

Let the earth's magnetic field be resolved into two components, H_z along the axis of z and H_x perpendicular to it, and let the axis of x be parallel to, and in the same sense as, the component H_x.

The magnetic field of the wave itself being negligible, the equations of motion of an ion are

$$m x = e\mathrm{P} + e\mathrm{H}_z y,$$
$$m y = e\mathrm{Q} - e\mathrm{H}_z x + e\mathrm{H}_x z,$$
$$m z = e\mathrm{R} - e\mathrm{H}_x y.$$

The electrodynamical circuital equations are, with Maxwell's notation,

$$\mathrm{curl}\,(\alpha, \beta, \gamma) = c^{-2}(\dot{\mathrm{P}}, \dot{\mathrm{Q}}, \mathrm{R}) + 4\pi \mathrm{N}e\,(x, y, z),$$
$$\mathrm{curl}\,(\mathrm{P}, \mathrm{Q}, \mathrm{R}) = -(\dot{\alpha}, \dot{\beta}, \dot{\gamma}),$$

where the specific inductive capacity of the medium has been taken as unity.

Assume a factor $e^{ip(t+z/c')}$ and let

$$\mu = c/c', \quad \mathrm{L} = 1 - \mu^2, \quad \mathrm{H}^2 = \mathrm{H}_x^2 + \mathrm{H}_z^2,$$

$$p_1 = -\frac{e}{m}\mathrm{H}, \quad p_2 = -\frac{e}{m}\mathrm{H}_x, \quad p_3 = -\frac{e}{m}\mathrm{H}_z,$$

so that $p_1^2 = p_2^2 + p_3^2$, and p_1, p_2, p_3 are positive when H_z and H_x are positive. (The effective ions are taken to be electrons,* so that e is negative.)

Also, let $\tilde{\omega}^2 = 4\pi N e^2 c^2/m$.

Then

$$\gamma = 0, \qquad \beta = -P/c', \qquad \alpha = Q/c',$$

$$x = -\frac{e}{m}\frac{L}{\tilde{\omega}^2}P, \qquad y = -\frac{e}{m}\frac{L}{\tilde{\omega}^2}Q,$$

$$R = -\frac{Lpp_2}{p^2-\tilde{\omega}^2}iQ, \qquad z = -\frac{e}{m}\frac{R}{\tilde{\omega}^2},$$

$$\left(1-\frac{\tilde{\omega}^2}{Lp^2}\right)P - \frac{ip_3}{p}Q = 0, \quad \left(1-\frac{\tilde{\omega}^2}{Lp^2}-\frac{p_2^2}{p^2-\tilde{\omega}^2}\right)Q + \frac{ip_3}{p}P = 0.$$

Eliminate P/Q from the last two equations. Then, with ν for $1-\tilde{\omega}^2/Lp^2$,

$$\nu^2(p^2-\tilde{\omega}^2) - \nu p_2^2 - p_3^2(p^2-\tilde{\omega}^2)/p^2 = 0.$$

If ν_1, ν_2 are the two values of ν, and P_1, P_2, Q_1, Q_2 the corresponding values of P and Q,

$$P_1/iQ_1 = p_3/p\nu_1. \qquad P_2/iQ_2 = p_3/p\nu_2.$$

and, since $\nu_1\nu_2 = -p_3^2/p^2$,

$$P_1/iQ_1 = -iQ_2/P_2$$

Hence also

$$\beta_1/i\alpha_1 = -i\alpha_2/\beta_2.$$

Solving the quadratic for ν gives

$$\nu = 1-\frac{\tilde{\omega}^2}{Lp^2} = \frac{\frac{1}{2}p_2^2 \mp \sqrt{\{\frac{1}{4}p_2^4 + p_3^2(p^2-\tilde{\omega}^2)^2/p^2\}}}{p^2-\tilde{\omega}^2}.$$

Hence, after a little simplification,†

$$L = 1-\mu^2 = \frac{\tilde{\omega}^2}{p}\frac{p(p^2-\tilde{\omega}^2-\frac{1}{2}p_2^2) \mp \sqrt{\{\frac{1}{4}p^2p_2^4 + p_3^2(p^2-\tilde{\omega}^2)^2\}}}{p^4-p^2(p_1^2+\tilde{\omega}^2)+p_3^2\tilde{\omega}^2}, \quad \text{(A)}$$

where the index μ is c/c'.

In order that we should be able to distinguish between the two values of L, the square root is always to be taken positive. The quantity under the root cannot vanish, and the two values of L are always distinct, except in the special case $p_2 = 0$.*

---

* If both electrons and positive ions are present in about equal numbers, the latter, being much more massive, may be neglected.

† [*Added August* 16, 1928.—Since writing this paper I have learnt that this formula, with a different notation, was given in a paper by Prof. Appleton read to the Washington Radio Conference in October, 1927. *Vide* 'Papers of the General Assembly of 1927, Union Radio Scientifique Internationale.' It also occurs in a question set by Mr. E. Cunningham in the Mathematical Tripos in 1928.]

Hence, in general, the wave is split up into two elliptically polarised waves travelling with different velocities The magnetic vector lies in the wave front, but the electric vector is oblique to the wave front. The axes of the magnetic ellipse are along and perpendicular to the component, in the wave front, of the imposed magnetic field, one of the axes of the electric ellipse is in the direction of this component also The paths of the ions are ellipses in a plane oblique to the wave front. Usually L will be small, and the inclination of the electric vector to the wave front will be small; but the inclination to the wave front of the plane in which the paths of the ions lie will, in general, be appreciable. The polarisation will be right-handed† for the upper sign and left-handed for lower sign in (A), when, as here, the imposed magnetic field has a component in a direction opposite to that of propagation When the imposed magnetic field has a component in the direction of propagation, the polarisation is left-handed for the upper sign and right-handed for the lower sign. For the upper sign in (A) the minor axis of the magnetic ellipse and the major axis of the electric ellipse are in the direction of the component, in the wave front, of the imposed magnetic field, and conversely for the lower sign. The ratio of the axes of the magnetic ellipse is the same for each of the components into which the wave is split up Numerical examples of the polarisation of downcoming waves will be given later.

Special Case I.—Imposed Magnetic Field perpendicular to the Direction of Propagation.

(I).—The upper sign gives

$$L = \varpi^2/p^2, \qquad \mu^2 = 1 - \varpi^2/p^2,$$

as for no magnetic field.‡

In this case

$$\alpha = 0,\ \gamma = 0,\ Q = 0,\ R = 0,\ y = 0,\ z = 0,\ \beta = -P/c',\ x = -eP/mp^2.$$

The wave is plane polarised with the electric vector along, and the magnetic vector at right angles to, the imposed magnetic field. The induced motion of the electrons is along the imposed magnetic field. There is no component of electric force along the direction of propagation. The shorter the wave, the smaller the amplitude of the induced motion of the electrons.

---

* When specific inductive capacity K is introduced, the difference made in (A) is that ϖ^2/K must be written for ϖ^2, and L is $1-\mu^2/K$ instead of $1-\mu^2$.

† The polarisation is said to be right-handed if the direction of rotation of the magnetic vector appears right-handed when we look along the direction of propagation.

‡ The expressions for the refractive index in the two special cases were given by Appleton and Barnett, 'Proc. Camb. Phil. Soc.,' vol. 22, p. 674 (1925).

(II).—The lower sign gives

$$\mathrm{L} = \frac{\tilde{\omega}^2}{p^2}\,\frac{p^2 - \tilde{\omega}^2}{p^2 - p_1{}^2 - \tilde{\omega}^2}. \qquad \mu^2 = 1 - \frac{\tilde{\omega}^2}{p^2}\,\frac{p^2 - \tilde{\omega}^2}{p^2 - p_1{}^2 - \tilde{\omega}^2}.$$

In this case $\beta = 0,\ \gamma = 0,\ \mathrm{P} = 0,\ x = 0,\ \alpha = \mathrm{Q}/c',\ y = -\frac{e\mathrm{Q}}{mp^2}\,\frac{p^2 - \tilde{\omega}^2}{p^2 - p_1{}^2 - \tilde{\omega}^2}$,

$$\mathrm{R} = -\frac{p'}{p}\,\frac{\tilde{\omega}^2}{p^2 - p_1{}^2 - \tilde{\omega}^2}\,\iota\mathrm{Q}, \quad z = -\frac{p'}{p}\,\frac{e/m}{p^2 - p_1{}^2 - \tilde{\omega}^2}\,\iota\mathrm{Q}.$$

The wave is plane polarised as regards the magnetic vector, which is along the imposed magnetic field. The electric vector is elliptically polarised in a plane through the direction of propagation, perpendicular to the imposed magnetic field. The induced motion of the electrons is elliptic, in the same plane.

Special Case II.—Imposed Magnetic Field along the Direction of Propagation.

In this case

$$\mathrm{L} = \frac{\tilde{\omega}^2}{p\,(p \pm p_1)}, \qquad \mu^2 = 1 - \frac{\tilde{\omega}^2}{p\,(p \pm p_1)}.$$

The upper and lower signs correspond to the upper and lower signs in the general case if $p^2 > \omega^2$, and conversely if $p^2 < \tilde{\omega}^2$.

In this case $\gamma = 0,\ \ \mathrm{R} = 0,\ \ z = 0,\ \ \alpha = \mathrm{Q}/c',\ \ \beta = -\mathrm{P}/c'$,

$$\alpha = \mp\,\iota\beta,\ \ \mathrm{P} = \mp\,\iota\mathrm{Q},\ \ x = -\frac{e}{m}\,\frac{\mathrm{P}}{p\,(p \pm p_1)}, \qquad y = -\frac{e}{m}\,\frac{\mathrm{Q}}{p\,(p \pm p_1)}.$$

The wave is circularly polarised, for the imposed magnetic field in the direction opposite to that of propagation the polarisation is right-handed for the upper sign and left-handed for the lower sign; for the imposed magnetic field in the direction of propagation, the polarisation is right-handed for the lower sign and left-handed for the upper sign.

There is no component of electric force along the direction of propagation. The imposed motion of the electrons is circular, and is in the same phase as the electric force for the upper sign, and the same phase or directly opposite for the lower sign according as $p \gtrless p_1$.

The "Forbidden" Wave-lengths.

For certain values of p^2 and $\bar{\omega}^2$, μ^2 may be negative. In such cases the wave will not travel. The changes from permissible values to values for which the wave will not travel occur when $\mu^2 = 0$ or $\pm\infty$, *i.e.*, when L, equal to $1 - \mu^2$, is 1 or $\pm\infty$.

In the general case, with the imposed magnetic field oblique to the direction of propagation, $L = 1$ when

$$p^2 - \bar{\omega}^2 = 0,$$

or when

$$p^2 + pp_1 - \bar{\omega}^2 = 0,$$

or when

$$p^2 - pp_1 - \bar{\omega}^2 = 0.^*$$

Of these equations, the first refers to the upper sign in equation (A), and the other two to the lower sign.

We may note here that when $p^2 - \bar{\omega}^2 = 0$, the value of L for the lower sign is zero.

$L = \pm\infty$, when

$$p = 0,$$

or when

$$p^4 - p^2(\bar{\omega}^2 + p_1^2) + \bar{\omega}^2 p_3^2 = 0.$$

When $p = 0$, both values of L are infinite. For the upper sign, $L = -\infty$, $\mu^2 = +\infty$. For the lower sign, $L = +\infty$, $\mu^2 = -\infty$. When the second equation is satisfied, the value of L for the lower sign becomes infinite for $p > p_1$. If $p < p_3$, the value of L for the upper sign becomes infinite. If $p_1 > p > p_3$, neither value of L becomes infinite, since we are restricted to positive values of $\bar{\omega}^2$ and p^2.

When $p^2 \rightarrow \infty$, then $L \rightarrow 0$ and $\mu^2 \rightarrow 1$.

It is now easy to sketch the graphs of μ^2 against p for both the upper and lower signs in equation (A); the forbidden values of p are, then, those for which μ^2 is negative. In fig. 1, μ^2 is plotted against $p/10^6$ for $\bar{\omega}^2$, which is $4\pi Ne^2c^2/m$, equal to $9 \,.\, 10^6$, p_1 equal to $9{\cdot}7 \,.\, 10^6$, p_2 equal to $9{\cdot}2 \,.\, 10^6$ and p_3 equal to $3 \,.\, 10^6$. The general features, as regards zeros and infinities and limiting values for large and small p, are the same for any value of $\bar{\omega}^2$ and for any imposed magnetic field oblique to the direction of propagation. The positions of the zeros and infinities are found by solving equations given above; all that we shall note here are approximate values for the two cases, $\bar{\omega}/p_1$ small and $\bar{\omega}/p_1$ large. By finding approximate solutions of the equations, we find that for

---

* These results are most easily obtained by putting $\nu = 1 - \varpi^2/p^2$ in the quadratic for ν, and factorising the left-hand side.

$\tilde{\omega}/p_1$ small, μ^2 for the upper sign vanishes when $p = \tilde{\omega}$, and becomes infinite when $p = \tilde{\omega}p_3/p_1$; while μ^2 for the lower sign vanishes when $p = \tilde{\omega}^2/p_1$ and when $p = p_1 + \tilde{\omega}^2/p_1$, and becomes infinite when $p = p_1 + \tilde{\omega}^2 p_2^2/2p_1^3$. For $\tilde{\omega}/p_1$ large, μ^2 for the upper sign vanishes at $p = \tilde{\omega}$, and becomes infinite at $p = p_2$; while μ^2 for the lower sign vanishes at $p = \tilde{\omega} - \frac{1}{2}p_1$ and at $p = \tilde{\omega} + \frac{1}{2}p_1$, and becomes infinite at $p = \tilde{\omega} + p_2^2/2\tilde{\omega}$.

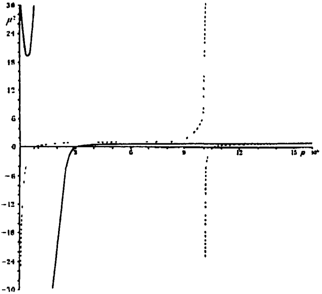


Fig. 1.—The Variation of μ^2 with p, where the index μ is c/c', and $p/2\pi$ is the frequency.

The wave-train to which the continuous line refers is polarised with a left-handed rotation about the direction of propagation when there is a component of imposed magnetic force *along* that direction. The wave-train to which the broken line refers is then polarised with a right-handed rotation. When the imposed magnetic field has a component in the direction *opposite* to that of propagation, the wave-train to which the continuous line refers is right-handedly polarised, and that to which the broken line refers is left-handedly polarised. The curves are drawn to scale for $4\pi Ne^2c^2/m$ equal to 9.10^{12}, so that N is 2812. The average value of the earth's magnetic field for England has been taken, with the wave-train travelling horizontally along a magnetic meridian, so that the horizontal component of the earth's field is along the direction of propagation (or along the direction opposite to it), and the vertical component is at right angles. The general features are the same for any value of N and for any imposed magnetic field oblique to the direction of propagation.

Hence, for $\tilde{\omega}/p_1$ small, the forbidden values of p are approximately

$$\tilde{\omega}p_3/p_1 < p < \tilde{\omega}$$

for the upper sign, and

$$p < \tilde{\omega}^2/p_1, \qquad p_1 + p_2{}^2\tilde{\omega}^2/2p_1{}^3 < p < p_1 + \tilde{\omega}^2/p_1$$

for the lower sign.

For $\tilde{\omega}/p_1$ large, the forbidden values of p are approximately

$$p_3 < p < \tilde{\omega}$$

for the upper sign, and

$$p < \tilde{\omega} - \tfrac{1}{2}p_1, \qquad \tilde{\omega} + p_2{}^2/2\tilde{\omega} < p < \tilde{\omega} + \tfrac{1}{2}p_1$$

for the lower sign.

Approximately, we may say that for $\tilde{\omega}/p_1$ large,

$$p < \tilde{\omega}$$

is forbidden for both components, except that $p < p_3$ is permissible for the one corresponding to the upper sign, since the short interval

$$\tilde{\omega} - \tfrac{1}{2}p_1 < p < \tilde{\omega} + p_2{}^2/2\tilde{\omega},$$

which is permissible for the lower sign, will be quite useless in practice.

Now the magnetic field of the earth is derived from a potential approximately given by

$$\Omega = 0\cdot32\, a^3 r^{-2} \cos\gamma,$$

where a is the radius of the earth, and γ the angular distance from a hypothetical magnetic north pole.

(Since the ionised region probably extends upwards from a height of the order of 50 kilometres, the local irregularities in the earth's magnetic field will be to some extent smoothed out.)

At a height of 50 kilometres, the horizontal and vertical components are given by

$$\mathrm{H}_h = 0\cdot31 \sin\gamma, \qquad \mathrm{H}_v = 0\cdot62 \cos\gamma.$$

When the wave is going into the ionised region inclined at an angle θ to the direction of the field,

$$\mathrm{H}_z = -\mathrm{H}\cos\theta, \qquad \mathrm{H}_x = \mathrm{H}\sin\theta.$$

When long distance transmission is taking place, and the ray is horizontal and makes an angle θ with the magnetic meridian,

$$\mathrm{H}_z = -\mathrm{H}_h\cos\theta, \qquad \mathrm{H}_x = (\mathrm{H}_h{}^2\sin^2\theta + \mathrm{H}_v{}^2)^{\frac{1}{2}}.$$

The total field H varies from horizontal and equal to 0·31 at the magnetic equator to vertical and equal to 0·62 at the poles.

Average values for England are 0·52 for the vertical component and 0·17 for the horizontal component. With electrons for the effective ions, so that $e/m = 1\cdot77 \,.\, 10^7$, approximate values of p_1, p_2, p_3 are given for three cases below.

$H_z = 0\cdot31$, $H_x = 0$; $p_1 = p_3 = 0\cdot55 \; 10^7$, $p_2 = 0$.

$H_z = 0$, $H_x = 0\cdot62$; $p_1 = p_2 = 1\cdot1 \,.\, 10^7$, $p_3 = 0$.

$H_z = 0\cdot17$, $H_x = 0\cdot52$; $p_1 = 0\cdot97 \cdot 10^7$, $p_2 = 0\cdot92 \,.\, 10^7$, $p_3 = 0\cdot3 \,.\, 10^7$.

Thus in the first case p_1 corresponds to a wave-length* of about 340 metres; in the second to 170 metres, and in the last to about 200 metres.

Also, with $e/m = 1\cdot77 \,.\, 10^7$ and $e = 1\cdot6 \,.\, 10^{-20}$, $\tilde{\omega}^2$, which $= 4\pi N e^2 c^2/m$, is $3\cdot2 \,.\, 10^9$ N approximately.

Some numerical examples of forbidden wave-lengths will now be given. We take the average value of the magnetic field for England, so that p_1 is approximately 10^7, and consider propagation along the magnetic meridian, so that $p_3 = 0\cdot3 \,.\, 10^7$. Then for $N = 300$, only the component corresponding to the lower sign will travel for wave-lengths between 2 and 6 kilometres; and so long as N does not exceed, say, $3 \,.\, 10^3$, both the limits of this interval are inversely proportional to $N^{\frac{1}{2}}$. In the same way, for $N = 300$, only the component corresponding to the upper sign will travel for wave-lengths greater than 19 kilometres, and the lower limit of this interval is inversely proportional to N so long as N does not exceed $3 \,.\, 10^3$. In addition, for wave-lengths about 200 metres, only the component corresponding to the upper sign will travel.

For $N = 3 \,.\, 10^5$, neither component travels for wave-lengths exceeding 60 metres; and the lower limit of this interval is inversely proportional to $N^{\frac{1}{2}}$, so long as N exceeds 10^5, say. An exception is that for N greater than 10^5, say, the component corresponding to the upper sign travels for wave-lengths greater than about 630 metres.

Another method of looking at the matter is to consider the variation of refractive index with N, for a given wave-length.

The Variation of Refractive Index with N *for a given Wave-length.*

The values of $\tilde{\omega}^2$ for which μ^2 vanishes or becomes infinite are at once obtained from equations given above.

When $\tilde{\omega}^2 = 0$, both values of μ^2 are 1. Also $\mu^2 = 1$ for the lower sign when $\tilde{\omega}^2 = p^2$.

---

* The term wave-length is used throughout in the sense of standard wave-length, namely $2\pi c/p$, where c is the velocity of light.

When $\bar{\omega}^2 = \infty$, $\mu^2 = \pm\infty$ according as $p \lesseqgtr p_2$ for the upper sign, and $\mu^2 = -\infty$ for the lower sign.

The graphs of μ^2 against N, equal to $\bar{\omega}^2/3 \cdot 2 \,.\, 10^9$, can now easily be sketched for the various cases that arise. The graphs have a different character, in the case of the upper sign, according as p is greater or less than p_3, and, in the case of the lower sign, according as p is greater or less than p_1. In figs. 2 and 3, μ^2 is plotted against N/10³ and N/10⁵ respectively. In both, p_1 has been taken as $9 \cdot 7 \,.\, 10^6$, p_2 as $9 \cdot 2 \,.\, 10^6$, and p_3 as $3 \,.\, 10^6$. In fig. 2, p has been taken

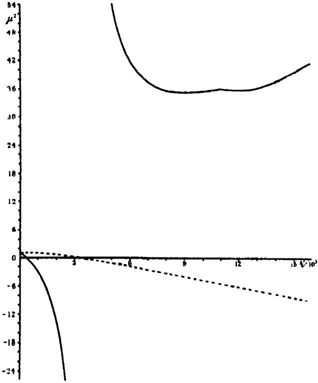


FIG. 2.—THE VARIATION OF μ^2 WITH N FOR LONG WAVES, WHERE N IS THE NUMBER OR FREE ELECTRONS PER CUBIC CENTIMETRE.

The polarisations of the wave-trains to which the continuous and broken lines refer are the same as in fig. 1. The curves are drawn to scale for $p = 10^6$, where $p/2\pi$ is the frequency. The corresponding wave-length is 1884 metres. The same intensity and direction of the magnetic field have been taken as in fig. 1. The general features are the same for any imposed magnetic field oblique to the direction of propagation, and for any wave-length, so long as p is less than p_3 for the continuous line curve, and less than p_1 for the broken line curve, where p_1 denotes $|eH/m|$, and p_3 denotes $|eH_z/m|$. H is the intensity of the imposed magnetic field, and H_z its component in the direction of propagation.

as 10^6, corresponding to a wave-length of 1884 metres. In fig. 3, p has been taken as $2 \,.\, 10^7$, corresponding to a wave-length of 94 metres. For any imposed magnetic field oblique to the direction of propagation, and for any wave-length, the general features for the upper sign are the same as for the continuous line curve in fig. 2 or fig. 3, according as p is less or greater than p_3; while for the lower sign the general features are the same as for the broken line curve in fig. 2 or fig. 3, according as p is less or greater than p_1. By solving equations given above, we easily find that for p greater than p_3, μ^2 for the upper sign vanishes when $\bar{\omega}^2 = p^2$; for p less than p_3, μ^2 for the upper sign vanishes when $\bar{\omega}^2 = p^2$, and becomes infinite when $\omega^2 = p^2(p_1^2 - p^2)/(p_3^2 - p^2)$. For p greater than p_1, μ^2 for the lower sign vanishes when $\bar{\omega}^2 = p^2 - pp_1$ and when $\bar{\omega}^2 = p^2 + pp_1$, and becomes infinite when

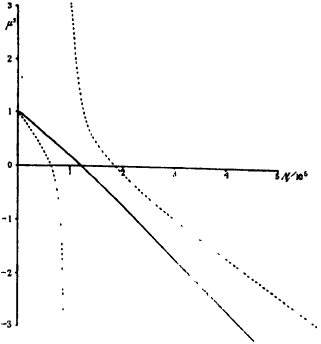


FIG 3.—THE VARIATION OF μ^2 WITH N FOR SHORT WAVES.

The polarisations of the wave-trains to which the continuous and broken lines refer are the same as in figs. 1 and 2. The curves are drawn to scale for $p = 2 \cdot 10^7$. The corresponding wave-length is 94 metres The same intensity and direction of the magnetic field have been taken as in figs 1 and 2. The general features are the same for any imposed magnetic field oblique to the direction of propagation, and for any wave-length, so long as p is greater than p_3 for the continuous line curve, and greater than p_1 for the broken line curve.

$\tilde{\omega}^2 = p^2(p^2 - p_1^2)/(p^2 - p_2^2)$; for p less than p_1, μ^2 for the lower sign vanishes when $\tilde{\omega}^2 = p^2 + pp_1$.

Again μ^2 must be positive. Further, in such a case as the continuous line curve of fig. 2 (upper sign, $p < p_2$), where μ^2 is positive for N sufficiently large, but negative for some smaller value of N, the ray will not be able to penetrate the layer in which μ^2 is negative. (If the distribution of N varies discontinuously, the ray may enter a region of N large without passing through all smaller values of N; this possibility will be left aside in the following discussion. The modifications to be introduced are sufficiently obvious.)

Thus we have the following restrictions on ω^2, which is $3 \cdot 2 \,.\, 10^9$ N.

For $p > p_1$, both waves will travel only if $\tilde{\omega}^2 < p^2 - pp_1$, while no waves will travel if $\tilde{\omega}^2 > p^2$. For intermediate values of $\tilde{\omega}^2$ the wave corresponding to the upper sign may travel.

For $p < p_1$, both waves will travel only if $\omega^2 < p^2$; no waves travel if $\omega^2 > p(p + p_1)$, for intermediate values of ω^2, the wave corresponding to the lower sign may travel.

Thus, with p_1 equal to $9 \cdot 7 \,.\, 10^6$, for a wave of 10 kilometres, neither component can penetrate to a layer in which N exceeds 580, while only one component will reach a layer in which N exceeds 11.* For a wave-length of 1 kilometre these limits of N are $6 \cdot 8 \,.\, 10^3$ and $11 \,.\, 10^2$ respectively; for a 400 metre wave, $21 \,.\, 10^3$ and $6 \cdot 8 \,.\, 10^3$ respectively; for 100-metre wave, $11 \,.\, 10^4$ and $5 \cdot 4 \,.\, 10^4$ respectively; while for a 20-metre wave they are $2 \cdot 8 \,.\, 10^6$ and $2 \cdot 5 \,.\, 10^6$.

Another noteworthy result is that for $p < p_1$, in the case of the lower sign, μ^2 increases with N at first. (In fig. 2 the scale is too small for this to be seen.) When the wave-length and the intensity and direction of the imposed magnetic field are known it is not difficult to calculate the maximum value of the refractive index and the value of N for which it occurs; a simple formula can be given for long waves for which p/p_1 is fairly small. For then

$$L = -\tilde{\omega}^2(p^2 - \tilde{\omega}^2)/p^2 p_2^2$$

approximately, the maximum of μ^2 occurs when $\tilde{\omega}^2 = \frac{1}{2}p^2$, and is $1 + \frac{1}{4}p^2/p_2^2$. In any case the maximum must occur for a value of $\tilde{\omega}^2$ less than p^2, except for the special case $p_2 = 0$, when the imposed magnetic field is along the direction of propagation, in which case μ^2 is increasing for all N.

---

* In the absence of the magnetic field both numbers would be 11. This shows how considerable may be the influence of the magnetic field for long waves.

The Polarisation of the Downcoming Wave.

The polarisation of the downcoming wave may be expected to be that with which it leaves the ionised region. Near the bottom of the ionised region, $\bar{\omega}^2$ will be small, so that the component of electric force along the direction of propagation will be quite negligible. Also for the upper sign, when $\bar{\omega}^2$ is small,

$$\nu = 1 - \frac{\bar{\omega}^2}{\mathrm{L}p^2} = -\frac{(\frac{1}{4}p_2{}^4 + p^2p_3{}^2)^{\frac{1}{2}} - \frac{1}{2}p_2{}^2}{p^2},$$

and $\nu\mathrm{P} - ip_3\mathrm{Q}/p = 0$, so that $\beta\nu + ip_3\alpha/p = 0$.

Particular interest attaches, as we shall see, to a wave-length of 400 metres; some results are given here for this wave-length and will be used later.

Let the ray make an angle θ with the direction of the magnetic field; and let $p = \frac{1}{2} \cdot 10^7$, $p_1 = 10^7$, $p_2 = 10^7 \sin\theta$, $p_3 = -10^7 \cos\theta$. Then, for the upper sign

$\beta = -i\alpha$ if $\theta = 0$;

$\beta = -i\alpha$ nearly if $\theta = 9°$;

$\beta = -1\cdot4\, i\alpha$ nearly if $\theta = 33°$;

$\beta = -1\cdot9\, i\alpha$ nearly if $\theta = 45°$.

The manner in which the polarisation depends on the wave-length may be illustrated here for $\theta = 30°$. Then for a 100-metre wave, $\beta = -1\cdot1\, i\alpha$; for a kilometre wave, $\beta = -2\cdot0\, i\alpha$; for a 10-kilometre wave, $\beta = -14\cdot4\, i\alpha$.

Since it has been shown that the magnetic ellipse for the second component wave-train is at right-angles to, and has its axes in the same ratio as, the magnetic ellipse for the first component, the rotation being opposite in the two cases, the results for the second component can at once be written down.

The Polarisation Measurements of Appleton and Ratcliffe.

Now Appleton and Ratcliffe* have measured the polarisation of the downcoming wave. The wave-length used was approximately 400 metres. Transmission took place from Birmingham and from the National Physical Laboratory, the receiver being at Peterborough. The components of magnetic force were measured in, and at right-angles to, the plane of propagation, which is a vertical plane through the sending and receiving stations. These components are denoted by $\mathrm{H}_1{}'$ and H_1, and are measured along axes that will be denoted here by 1 and 2. The angle of incidence of the downcoming wave varied between 24° and 30°; and for both the south to north and west to east transmission, the downcoming wave was approximately circularly polarised with a left-handed rotation.

---

* 'Roy. Soc. Proc.,' A, vol. 117, p. 576 (1928).

Appleton and Ratcliffe connect the circular polarisation with the fact that the downcoming ray is approximately along the earth's magnetic field; and the absence of the right-handed polarised wave is ascribed to increased dissipation of energy by collisional friction.

We proceed to consider in detail the theory and calculations for the wave-length and incidence in these experiments. A wave, plane polarised when it leaves the transmitter, with the electric vector in the plane of propagation, is split up on entering the ionised region into two component elliptically polarised waves, in a manner that can be calculated. Each component wave has its direction of propagation gradually bent round, and as this direction changes so does the state of polarisation. The two component waves recombine at the receiver. But the two component waves have travelled with different speeds in different paths, and have been damped by collisional friction to different extents; the phase difference and the ratio of the amplitudes are not the same on emerging as they were on entering. If these two quantities were known, the state of polarisation of the wave at the receiver could be completely described; conversely, these two quantities can be found if the polarisation at the receiver is known from experimental sources.

To determine the polarisation of the components it is necessary to know the inclination of the ray to the earth's magnetic field; to compare the theoretical with the experimental results it is necessary to know the inclination of the axes 1 and 2 to our axes of x and y. We proceed to find formulæ for these angles for the west to east transmission. The ray is assumed to be in the plane of propagation. Let i_0 and d be the angles the downcoming ray and the earth's magnetic field make with the downward vertical. Then the inclination of the ray to the magnetic field is easily found to be $\cos^{-1}(\cos i_0 \cos d)$. With $\cos i_0 = 0 \cdot 89$, which is its mean experimental value, and $0 \cdot 17$, $0 \cdot 52$ for the horizontal and vertical components of the magnetic field, the inclination is 33°.

Let the earth's field have components H_h horizontally to the north, and H_v vertically downwards. The downcoming ray is perpendicular to the former component, and makes an angle i_0 with the latter. Hence the component of the earth's field along the direction of propagation is $H_v \cos i_0$; in the wave front there are components H_h horizontally to the north, and $H_v \sin i_0$ in the plane of propagation. Thus the axis of x, which is along the resultant of the last two components, makes an angle whose tangent is $H_h/(H_v \sin i_0)$ with the negative direction of the axis 1. With $i_0 = 27°$, $H_h = 0 \cdot 17$, $H_v = 0 \cdot 52$, this gives an angle of 36°. The directions of 1 and 2 in fig. 4 have been chosen to correspond with those used by Appleton and Ratcliffe.

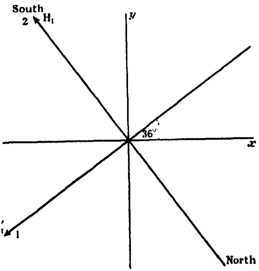


FIG. 4.—THE RELATIVE POSITIONS OF THE AXES OF x AND y AND THE AXES 1 AND 2 USED BY APPLETON AND RATCLIFFE.

The figure is drawn for a transmission at right angles to the magnetic meridian and for the downcoming ray at an angle of 27° with the vertical.

Since the inclination of the upgoing ray to the earth's magnetic field is 180°–33°, the wave on entering the ionised region is split up into two elliptically polarised components for one of which $\beta = 1 \cdot 4\, i\alpha$, while for the other $\beta = -i\alpha/1 \cdot 4$. The polarisation of each gradually changes*; for the first $\alpha = 0$, and for the second $\beta = 0$ when the ray is horizontal; when the ray emerges from the ionised region, the states of polarisation are represented by $\beta = -1 \cdot 4\, i\alpha$, $\beta = i\alpha/1 \cdot 4$.

Let us take $\alpha_0 \cos pt$, $-1 \cdot 4\alpha_0 \sin pt$ as the components of magnetic force along x and y for the first component of the incident wave; and $1 \cdot 4\, k\alpha_0 \cos (pt + \theta)$, $k\alpha_0 \sin (pt + \theta)$ as the corresponding quantities for the second component of the incident wave. Expressing the condition that the wave is plane polarised with magnetic vector along 2, we easily find $k^2 = 0 \cdot 8$ and $\theta = -107°$.

---

* Since the paths in the ionised region are different, the first and second components at the receiver will come from waves having slightly different angles of incidence. But this difference will usually be small. See, however, fig. 5 and the last paragraph on p. 278.

Hence $1/(1+k^2)$ or $0 \cdot 55$ of the energy of the incident beam goes into the first component, and the second component is 107° behind the first in phase.

The emergent beam may be represented, as regards its first component, by $-\alpha_1 \cos pt$, $-1 \cdot 4\, \alpha_1 \sin pt$, and as regards its second component by $1 \cdot 4\, \lambda\alpha_1 \cos (pt + \theta')$, $-\lambda\alpha_1 \sin (pt + \theta')$. (In writing down these expressions, use has been made of the fact that for the first component wave α vanishes, and for the second component wave β vanishes, when the ray is horizontal.) The amplitude of the first component has been reduced in the ratio α_1/α_0; that of the second component in the ratio $\lambda\alpha_1/k\alpha_0$. We shall call k/λ the relative damping factor and $\theta' - \theta$ the phase shift.

The experimental results show that the relative damping is considerable.* Since the phase shift will change continually during the night as the ionisation changes, $\theta' - \theta$ may be considered completely arbitrary. Thus the mean of the experimental results should agree with the calculated result for the first component wave only; while the irregularity of the experimental observations is to be attributed mainly to the presence of the second component wave, fairly heavily damped, with arbitrary and constantly changing phase difference. For the first component wave only, the components of magnetic force along 1 and 2 are $1 \cdot 15\, \alpha_1 \cos (pt - 45°)$ and $1 \cdot 27\, \alpha_1 \cos (pt + 62°)$, when the components along x and y are $-\alpha_1 \cos pt$, $-1 \cdot 4\, \alpha_1 \sin pt$, as above. The ratio of the amplitudes is $1 \cdot 1$ and the phase difference 107°. These should be compared with the experimental results.

Next, for accurate circular polarisation, we must have $\theta' = \pm 180°$, $\lambda = 1/6$. This gives $-77°$ for the phase shift and $5 \cdot 4$ for the relative damping factor. If we take the same relative damping factor, but $\theta' = 0$, we find that the components of magnetic force along 1 and 2 are $1 \cdot 14\, \alpha_1 \cos (pt - 57°)$ and $1 \cdot 39\, \alpha_1 \cos (pt + 71°)$, which give an amplitude ratio of $1 \cdot 2$ and a phase difference of 128°. This is just about within the limits of the observed departure from circular polarisation. Thus we may say that the experimental observations demand a relative damping factor of not less than about 5·4 or so.* Now, as Appleton and Ratcliffe point out, for propagation parallel to the imposed field, and for the wave-length used, the damping by collisional friction is about nine times as much, owing to the greater velocity of the electrons, for the second component wave as for the first. For other directions of propagation the factor is less, but depends on N; from rough calculations it appears that 6 or 7 is a probable average. If the first component has its amplitude reduced in the ratio $1 : \sigma$, and the second its amplitude reduced in the ratio $1 : \sigma^6$, the relative damping factor is σ^5. If this is 5·4, σ must be 1·40.

---

* An explanation for the absence of the second component wave might appear to be that the 70 or 80 miles between transmitter and receiver is greater than the skip distance for the first component, if this skip distance exist; but less than the skip distance for the second component. This explanation must at once be rejected, for it leads to a maximum value of N lying between 10^4 and $2 \, . \, 10^4$. With $2 \, . \, 10^4$ as the maximum value of N communication would be practically impossible, owing to the inevitable escape of the waves into space, with waves shorter than 36 metres, even when the bottom of the reflecting layer is taken no higher than 64 kilometres, or one-hundredth of the earth's radius, above the surface of the earth.

The calculation can similarly be carried out in detail for the south to north transmission. With the data previously used, the upgoing and downcoming rays make angles of 45° and 9° respectively with the earth's magnetic field. For the upgoing ray the axis of x coincides with the axis 1, and the axis of y is in the negative direction of 2; for the downcoming ray, x is in the negative direction of 1, and y coincides with 2. The upgoing wave is split up on entering the ionised region into two elliptically polarised waves, with components of magnetic force along x and y represented by $\alpha_0 \cos pt$, $-1 \cdot 9\, \alpha_0 \sin pt$ for the first component, and $1 \cdot 9\, k\alpha_0 \cos (pt + \theta)$, $k\alpha_0 \sin (pt + \theta)$ for the second component. The condition that the incident wave is plane polarised with magnetic vector along 2 gives $\theta = 180°$, $k = 1/1 \cdot 9$. On emerging, both components are practically circularly polarised. With a relative damping factor of 10/1·9 or 5·3, which is about the same as that arrived at before, the calculated results fit very readily within the experimental observations, and a relative damping factor of about 4 may be assumed without much discrepancy. This determination can only be very rough, because of variations in ionisation and experimental errors of observation. However, with a relative damping factor of 4, a reduction of the amplitude of the first component wave in the ratio 1 : 1·3 is likely

We may briefly note here the result to be expected for a north to south transmission with the same wave-length and incidence. As already pointed out, the mean of the experimental observations should agree with the result calculated for the first component only. This gives a phase difference of 90° and an amplitude ratio of 1·9. The incident wave is split up into two nearly circularly polarised components, with equal energy in each; and, with a relative damping factor of about 5, rather large variations are to be expected in the experimental results. Thus with a phase difference of 90°, the amplitude ratio may vary from 1·3 to 3·4.

---

* It is of interest to observe that even were the relative damping factor 1, the two component waves would not combine to give a plane polarised wave on emerging. With no phase shift, the magnetic force would have components $0 \cdot 68\, \alpha_1 \cos (pt + 27°)$ and $2 \cdot 22\, \alpha_1 \cos (pt + 47°)$ along 1 and 2; with a phase shift of 180°, these would be $2 \cdot 22\, \alpha_1 \cos (pt - 63°)$ and $0 \cdot 71\, \alpha_1 \cos (pt + 118°)$ respectively.

On the Distribution of Ionisation and the absence of the Right-handed Polarised Wave.

The main difficulty in explaining the results of Appleton and Ratcliffe appears to be that if the path were anything like that usually pictured,* with the ionised region beginning at a height of 90 to 110 kilometres, damping by collisional friction would hardly produce any appreciable result at all in such a short path. Yet it seems impossible to account for the experimental results without appreciable damping.

It is necessary to mention here a possible hypothesis. Appleton and Ratcliffe record† that, sometimes, in the middle of a winter night, when determining the height of the region by means of measurements of t_0, a height of about 250 kilometres would be found. Correspondingly great heights would be obtained by counting the number of "fringes" due to a quick wave-length change at the transmitter, though the heights determined by these methods do not show a good agreement. Also, in the periods before and after these results were obtained, the "fringes" give clear evidence of two downcoming waves with widely different amplitudes and path differences. From this is to be inferred that an upper layer at a height of about 200 kilometres exists, that usually in the night the maximum number of ions per cubic centimetre in the lower layer is sufficient to retain the first component wave, but that the second component wave penetrates through to the upper layer, and is gradually reflected back from there. The first and second components at the receiver thus arise from waves that have a different angle of incidence. This is shown in fig. 5. Sometimes in the middle of a winter night, the maximum number of ions in the lower layer falls sufficiently low for both component waves to penetrate through to the upper layer.

This description appears to fit the facts. The possible hypothesis concerning the polarisation results is that the experiments were all performed when the first component wave was reflected from the lower layer, and the second from the upper layer. The comparatively small amplitude of the second component might perhaps then be accounted for, since (1) the second component comes from a wave which leaves the transmitter at a much smaller angle with the vertical, and therefore probably has much less energy, and (2)

---

* *E.g.*, Appleton and Barnett, 'Roy. Soc. Proc.,' A, vol. 109, p. 621 (1925).

† 'Roy. Soc. Proc.,' A. vol. p. 115, 303 (1927).

not only is the second component much more strongly damped than the first for the same path, but it also has a much longer path in the ionised region.

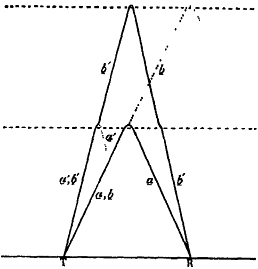


FIG. 5.—SHORT DISTANCE TRANSMISSION WITH A 400-METRE WAVE AT NIGHT.

T is the transmitter, R is the receiver. On entering the ionised region, each wave-train is split up into two component wave-trains, a and b The former, a, is reflected at a height of about 100 kilometres, but b penetrates through to a height of 200 kilometres or more before being reflected. Hence the two components, a and b', received at R, come from rays that leave the transmitter at quite different angles with the vertical.

Even if the polarisation experiments were performed when both components were reflected from the upper layer, the greatly increased length of path in the ionised region would perhaps give the different damping a chance to have an appreciable effect I had already performed some calculations with this hypothesis as basis, with results for the necessary relative damping very similar to those obtained above, when I learnt from Mr. Ratcliffe that it would have to be abandoned. Measurements of the field intensity during the period when both components were reflected from the upper layer had shown that this was as great and sometimes greater than the field intensity when one component was reflected at the lower layer. On examination, the transmitting antennæ were seen to have considerable horizontal portions. The experimental observation (hitherto unpublished) also gives us the important information that damping between the two layers is inconsiderable, so that the space between them is but sparsely populated with ions.

There remains, then, no alternative to attributing the weakness of the right-handed component to damping by collisional friction. To obtain any appreciable damping, one of two postulates is necessary. Either the critical wavelength, $2\pi c/p_1$, is much nearer to 400 metres than has so far been supposed; or the path in the ionised region must be much longer than has been supposed, and part of it must be low down where there is a fairly high pressure. The former postulate demands a large decrease in the earth's magnetic field: a decrease which must take place in the vertical component and reduce it, say, to half the value taken here. Although the currents induced in an ionised layer by its motion, under tidal forces, in the earth's magnetic field will cause a change in this field at the bottom of the ionised layer, it is extremely unlikely that this change will be considerable.* We are driven then to the second postulate: that the bottom of the ionised region is much lower than 100 kilometres, and that the number of ions per cubic centimetre up to about this height (100 kilometres), while not sufficient to cause considerable bending in a 400-metre wave, is sufficient to cause considerable damping under the pressure obtaining at these lower heights. We have thus arrived at the notion of discriminating between three layers of ionisation in the atmosphere.†

[*Added August* 16, 1928.—Then, whether it be during the day-time, when both components are reflected at the middle layer, or shortly after sunset or before sunrise, when one component is reflected at the middle layer, and one at the top layer, we should expect the measured polarisation to be predominantly left-handed for a 400-metre wave. The absence of the right-handed component will be due to increased dissipation by collisional friction, and dissipation in both cases will take place mainly in the bottom layer. In the middle of a winter night, both components are reflected from the top layer; again, and for the same reason as before, the measured polarisation might be expected to be predominantly left-handed; but recombination of ions will have taken place to a considerable extent in the bottom layer, where there is still a fairly considerable pressure, and so the damping will be less than before. The right-handed component should therefore be present in greater relative strength than just after sunset or before sunrise.

Some mention may be made of rays that have suffered multiple reflexions. A ray reflected once at the surface of the earth, and twice at the middle or top layer, will have passed four times through the bottom layer, and will be very considerably damped, except perhaps in the middle of the night, when recombination of ions has taken place in the bottom layer. Thus we should expect these multiply reflected rays to be more in evidence in the middle of the night than at other times.]

---

* Wireless experiments with a 200-metre wave should decide this point beyond doubt.

† *Cf.* Appleton, 'Nature,' vol. 100, p. 330 (1927), where a similar suggestion is made.

The experimental evidence for the existence of the two upper layers by Appleton and Ratcliffe has already been mentioned. There is further evidence for the existence of the topmost layer and the lowest layer. Thus the height deduced by T. L. Eckersley for long waves (over 10 kilometres) is about 50 kilometres.* Taylor and Hulburt† deduce a height of about 240 kilometres from skip distance data for short waves, while Breit and Tuve‡ also give heights of this order of magnitude. Generally speaking, we may say that experiments on long waves give heights from 40 to 70 kilometres; experiments on waves in the broadcasting band give heights of 90 to 110 kilometres, and experiments on short waves give heights of from 200 to 250 kilometres.

[*Added August* 16, 1928.—

On the Energy in the Downcoming Wave.

We have seen how to calculate the fraction of the energy of the incident wave that goes into each of its components; and we have found the minimum relative damping factor that will allow our calculated results to fit the experimental results. From the relative damping factor we can find how much each component is damped, and so we can find the ratio of the energy in the downcoming wave to the energy in the incident wave. Since we shall use the least value of the relative damping factor, we shall find an upper limit to this ratio. Again, the polarisation of each component is gradually changed as it changes its direction of propagation in the ionised region, and from this there results a further loss of energy. Thus our results will be definitely in excess.

In Appleton and Ratcliffe's experiments, we found for the west to east transmission that 0·55 of the energy goes into the first component wave, and 0·45 into the second. We also decided that the relative damping factor must sometimes have fallen as low as 5·4, and that the first component wave then has its amplitude reduced in the ratio 1 : 1·40, and the second in the ratio 1 : 7·56. Thus the energy in the downcoming wave, expressed as a fraction of that in the incident wave, is $0 \cdot 55/(1 \cdot 40)^2 + 0 \cdot 45/(7 \cdot 56)^2$, or $0 \cdot 281 + 0 \cdot 008$, or $0 \cdot 29$ nearly.

---

* Round, Eckersley, Tremellen and Lunnon, 'J. Inst. Elec. Eng.,' vol. 63, p. 933 (1925). Also T. L. Eckersley in the discussion on a paper by J. Hollingworth, 'J. Inst. Elec. Eng.,' vol. 64, p. 590 (1926).

† 'Phys. Rev.,' vol. 27, p. 189 (1926).

‡ 'Phys. Rev.,' vol. 28, p. 554 (1926).

For the south to north transmission, $1/(1 + k^2)$ of the energy goes into the first component, where k is $1/1 \cdot 9$. This gives $0 \cdot 78$ for the first component and $0 \cdot 22$ for the second. The relative damping factor sometimes fell as low as 4, so that the amplitude of the first component was then reduced in the ratio $1 : 1 \cdot 3$, and that of the second in the ratio $1 : 5 \cdot 2$. The energy in the downcoming wave, expressed as a fraction of that in the incident wave, is therefore $0 \cdot 78/(1 \cdot 3)^2 + 0 \cdot 22/(5 \cdot 2)^2$ or $0 \cdot 46 + 0 \cdot 008$, or $0 \cdot 47$.

As previously mentioned, these estimates will be in excess, but the ratio should in any case be greater for the south to north than for the west to east transmission, since a larger fraction of the energy goes into the first component wave, which is much less damped than the second.

From simultaneous measurements giving the angle of incidence, the polarisation, and the energy ratio, we could calculate exactly how much loss of energy is due to the change of " type " as the direction of propagation changes.]

On Wireless Communication and the Distribution of Ionisation.

The experimental evidence so far published is not sufficient to allow us to decide the facts with certainty, but a provisional picture can already be given which approximates to them with a considerable degree of probability, and which may serve as a basis for further discussion. The picture we have formed is that of three layers, and heights as given above. For short distance transmission with long waves (about 10 kilometres), the waves penetrate into the lowest layer, and are gradually refracted down to earth again. As the distance, and therefore the angle of incidence, increases, the penetration gets less and less; in long distance communication the waves hardly penetrate at all, and are reflected at almost glancing incidence. Waves in the broadcasting band are bent down to earth again in the second layer; short waves may penetrate right through to the topmost layer, travelling near the bottom of the layer in long distance communication, or even in the middle layer in the daytime, and penetrating up to great heights (250 kilometres or 50 kilometres into the upper layer would seem a fair estimate) for short distance communication.

If, as suggested here, short waves travel near the bottom of the layer in long distance communication, the waves most effective in establishing such communication are those leaving the transmitter at a small angle with the earth's surface; and the waves received should also come in at a small angle.

Whether this is true or not does not yet appear to be definitely decided, though there is some evidence* to show that it is so.

Damping by collisional friction takes place mainly in the lower layer; the paths of short waves and of medium waves being of about the same length in this region, the medium waves will be more damped, in proportion to the square of the wave-length. Long waves, on the other hand, may have a much shorter path in this region, and so be much less damped than the square of the wave-length law would imply.

The middle layer has a maximum of about 10^4 to $2 . 10^4$ electrons per cubic centimetre at night. Daytime values are not yet known with any certainty. No reliable estimate can yet be given for the lower layer, except that the number must be less than $4 . 10^3$, since 600-metre waves are known to penetrate to the second layer. An examination† of the latest skip distance data suggests a usual maximum content in the upper layer as great as 10^6 to $2 . 10^6$, the daytime value being somewhat more than twice the night value. It appears that the number may sometimes rise as high as 10^7.

Between the middle and upper layers the electron content is small compared with its value in, say, the middle layer; but between the middle and lower layers the electron content probably does not fall off much, if at all, from any maximum value there might be in the lower layer.

No information is yet available as to the gradients of ionisation.

It will clearly be necessary to account for the presence of these ionised regions in the atmosphere, and to fit in the picture here given with other evidence from terrestrial magnetism, auroræ, meteors, etc. No attempt will be made to do so here, but a few relevant facts may be noted. The bottom layer is clearly connected with the ozone layer; it is also roughly in the layer in which occurs the Lindemann–Dobson temperature inversion; and increase of temperature will help to explain the comparatively great absorption. The height of the middle layer coincides roughly with the auroral height; the spectra of auroræ contain bands due to molecular nitrogen in the singly ionised state, and a line due to oxygen mixed with helium, neon or argon. The main difficulty appears to be to account for the large electron content at great heights. If, as suggested by Chapman and Milne, hydrogen be absent from the upper atmosphere, then at 250 kilometres the atmosphere must be predominantly helium, even if allowance be made for temperature inversion and for the fact that more mixing takes place at heights of 100 kilometres or so than was previously supposed. It does not seem easy to account for the large ionisation of helium that is required by means of the known ionising agents. It is perhaps relevant that the amount of helium in the atmosphere appears to be far too little to account for the helium given off by radioactive rocks during the earth's history.*

---

* T. L. Eckersley, 'J. Inst. Elec. Eng.,' vol. 65, p. 600 (1927); Oscanyan, 'Proc Inst. R. Eng.,' vol. 15, p. 425 (1927).

† For a spherically stratified medium, $\mu(a+h)\sin\phi$ is constant along a ray, where μ is the refractive index, h is height above the earth's surface, and ϕ the angle of incidence of the ray on to a spherical layer. If either the angle, i_0, which the ray makes with the vertical at the earth's surface, or the height, h_0, at which the ray becomes parallel to the earth's surface, is known, the other can be roughly calculated from geometrical considerations. The value of μ, and hence the value of N, where the ray becomes parallel to the earth's surface can then be found from the formula above. For a rough estimate, the effect of the earth's magnetic field on the propagation of short waves may be neglected. See Appleton and Barnett, 'Roy. Soc. Proc.,' A, vol. 109, p. 621 (1925); Taylor and Hulburt, 'Phys. Rev.,' vol. 27, p. 189 (1926); Appleton, 'Proc. Camb. Phil. Soc.,' vol. 23, p. 155 (1926). For skip distance and similar data, reference may be made to Heising, Schelling and Southworth, 'Proc. Inst. R. Eng.,' vol. 14, p. 613 (1926); Hallborg, *ibid.*, vol. 15, p. 501 (1927); and Meissner, *ibid.*, vol. 15, p. 928 (1927).

The Polarisation of Very Short and Very Long Waves.

It is clear that the calculations already given for a 400-metre wave can be carried out in detail for any wave-length and any circumstances of propagation. Attention may be called to some general results for very short and very long waves.

So long as the upgoing ray makes only a moderate angle with the earth's magnetic field, short waves are split up into two almost circularly polarised waves. Since the path and the damping will not be very different for the two components, both waves will in general be received, and, if the downcoming rays make only a moderate angle with the earth's magnetic field, will combine to give a plane polarised wave, with direction varying rapidly with path difference.

For very long waves, the two components are almost plane polarised, even for moderate inclinations to the magnetic field. If only one component is received, it will in general (unless it comes down almost along the lines of magnetic force) be plane polarised; rapidly varying elliptic polarisation is to be expected if both components are received. But it should be noted that in this case the ionisation, and therefore also the refractive index, may undergo a considerable change in a very small fraction of a wave-length at the bottom of the ionised region, and may then not change much with height for a considerable distance, so that an appreciable reflected wave may be expected, especially at certain angles of incidence.

---

* Jeffreys, 'The Earth,' p. 254.

A main object of close discussion such as the present is to probe the adequacy of the current theory under new conditions. The very large electron content required for short distance transmission of short waves raises the question of the probability of some additional cause of bending; but it seems hard to imagine any other cause.

The author desires to thank Sir Joseph Larmor for the interest he has taken in the preparation of this paper.

The Thermal Conductivities of Carbon Monoxide and Nitrous Oxide.

By H. Gregory, Ph.D., A.R.C.S., D.I.C., Assistant Professor of Physics, and
C. T. Archer, M.Sc., A.R.C.S., D.I.C., Lecturer in Physics.

(Communicated by H. L. Callendar, F.R.S.—Received June 20, 1928.)

The authors' experiments on the thermal conductivities of carbon monoxide and nitrous oxide were undertaken partly because very few determinations had been made previously, and partly on account of a consideration of other physical properties of these gases. Smith* showed experimentally that the viscosities of nitrogen and carbon monoxide are equal, and a similar result was obtained in the case of carbon dioxide and nitrous oxide. Such results are indicated by the Kinetic Theory of Gases from the aspect of the equality of molecular weights in the two cases. Similar equalities are not anticipated, however, in the case of the thermal conductivities, as the conduction effect depends on a consideration of differences in molecular structure. The following table shows the values of the thermal conductivities and the viscosities of the four gases concerned, and illustrates the extent to which the thermal conductivities differ :—

—	CO.	N_2.	N_2O.	CO_2.
$\eta . 10^7$	1665	1665	1366	1366
$K . 10^7$	563·3	580	374	360·4

---

* 'Proc. Phys. Soc.,' vol. 34, p. 155 (1922).

With regard to previous work, the values of the thermal conductivities and the temperature coefficients of conductivity, where available, of CO and N_2O are shown below :—

—	$K_0 . 10^7$.		Temperature coefficient.	
	CO.	N_2O	CO.	N_2O.
Eucken	542·5	351·5		
Krey	525·0			
Moser	542·6			
Winkelmann	499·2	350·0	$45 . 10^{-4}$	$28·8 . 10^{-4}$
Stefan	551·8	355·5		
Weber, S.		353·0	$44 . 10^{-4}$	
Mean value	532·2	352·5	$44·5 . 10^{-4}$	

It should be pointed out that in the above table the results of Eucken and Moser* were obtained from relative observations with respect to air, using the value 566 . 10^{-7} at 0° C., and that of Stefan† was obtained also relatively to air at 10° C., using the value 581 . 10^{-7} at 10° C., this value being reduced to 0° C. by calculation, using the temperature coefficient for air obtained by the authors in their own determination.

Description of the apparatus used and account of the experimental procedure are not considered necessary here. The apparatus was identical with that used in the experimental investigation of the thermal conductivities of carbon dioxide, oxygen and nitrogen carried out by Gregory and Marshall,‡ and the experimental procedure is very fully described in the papers on that work and on the original work§ carried out by the authors of the present paper. Reference should be made to these papers for further details.

A brief description of the mode of preparation of the gases is given :—

Carbon Monoxide.—For the preparation of this gas, pure re-distilled concentrated sulphuric acid and pure dry sodium formate were used. The acid was allowed to drip slowly on to the sodium formate, the mixture being heated gently if necessary. Possible impurities were removed by repeated washings in a strong solution of pure caustic potash, and the gas dried by prolonged exposure to pure phosphorus pentoxide.

---

* Eucken, 'Phys. Z.,' vol. 14, p. 324 (1913); Moser, Diss. Berlin (1913).

† 'Wien. Akad. Ber.,' vol. 72 (1875).

‡ 'Roy Soc. Proc.,' A, vol. 118, p. 594 (1928).

§ 'Roy. Soc. Proc.,' A, vol. 110, p. 91 (1926).

Nitrous Oxide.—The gas was prepared from pure sodium nitrite and pure hydroxylamine hydrochloride, a saturated solution of the former being allowed to drip into a strong solution of the latter, gentle heat being applied to the mixture as required. The gas evolved was passed through a series of wash bottles containing a solution of ferrous sulphate, to which sulphuric acid had been added, in order to remove possible impurities such as nitric oxide. It was then passed through a second series of wash bottles containing pure concentrated sulphuric acid to remove moisture, and finally dried over specially prepared phosphorus pentoxide.

In both cases analysis showed that the gases used were free from impurities.

The following tables indicate the observations made and the results obtained.

The symbols used are as follows :—

θ = Temperature of heated wire.
$(R/l)_\theta$ = Resistance per unit length of wire, in ohms, at the temperature $\theta°$ C.
ϕ_θ = Correction for radiation at temperature $\theta°$ C.
θ_n' = Temperature drop through wall of narrow tubes.
θ_w' = Temperature drop through wall of wide tubes.
P = Pressure of gas in centimetres of mercury.
C_n = Current, in amperes, for narrow tubes.
C_w = Current, in amperes, for wide tubes.
K_n = Apparent conductivity of gas, narrow tubes.
K_w = Apparent conductivity of gas, wide tubes.
θ_m = Mean temperature of gas.
K_m = Conductivity of gas at temperature $\theta°_m$ C.
α = Value of constant obtained in determining K_m.

In all cases the bath temperature was 0° C., the tubes being immersed in a well-stirred mixture of ice and water.

Table I.—Carbon Monoxide.

$\theta = 9{\cdot}188°$ C. $(R/l)_\theta = 0{\cdot}12939$. $\phi_\theta = 0{\cdot}00000181$ cal./cm.

$\theta_n' = 0{\cdot}006°$ C. $\theta_w' = 0{\cdot}002°$ C.

P.	C_n.	K_n.	C_w.	K_w.
74·74	0·15018	0·00005733	0·13829	0·00005746
65·09	0·15018	5733	0·13820	5737
55·67	0·15013	5729	0·13815	5733
46·59	0·15013	5729	0·13809	5729
40·57	0·15008	5726	0·13804	5725
36·54	0·15008	5726	0·13804	5725
32·14	0·15003	5722	0·13798	5720
28·53	0·15003	5722	0·13794	5716
24·17	0·14998	5718	0·13794	5716
19·93	0·14998	5718	0·13789	5712
16·23	0·14993	5714	0·13789	5712
12·07	0·14983	5707	0·13778	5704
7·55	0·14970	5697	0·13768	5696
2·55	0·14878	5626	0·13693	5633

$\theta_m = 4{\cdot}596°$ C. $K_m = 0{\cdot}00005715$. $a = 1{\cdot}1836$.

Table II.—Carbon Monoxide.

$\theta = 15{\cdot}284°$ C. $(R/l)_\theta = 0{\cdot}13236$. $\phi_\theta = 0{\cdot}00000445$ cal./cm.

$\theta_n' = 0{\cdot}010°$ C. $\theta_w' = 0{\cdot}004°$ C.

P.	C_n.	K_n.	C_w.	K_w.
78·70	0·19257	0·00005789	0·17796	0·00005837
67·16	0·19253	5786	0·17742	5808
56·33	0·19247	5783	0·17716	5790
48·04	0·19247	5783	0·17690	5774
39·04	0·19243	5781	0·17685	5770
34·81	0·19243	5781	0·17685	5770
30·96	0·19237	5778	0·17680	5767
27·33	0·19232	5774	0·17680	5767
23·42	0·19228	5771	0·17675	5764
19·49	0·19222	5769	0·17670	5761
15·45	0·19217	5765	0·17665	5757
11·92	0·19213	5762	0·17660	5754
8·26	0·19192	5750	0·17641	5741
3·19	0·19103	5698	0·17571	5695

$\theta_m = 7{\cdot}646°$ C. $K_m = 0{\cdot}00005768$. $a = 1{\cdot}1839$.

Table III.—Carbon Monoxide.

$\theta = 21{\cdot}390^\circ$ C. $(R/l)_\theta = 0{\cdot}13533$. $\phi_\theta = 0{\cdot}00000712$ cal./cm.

$\theta_n' = 0{\cdot}015^\circ$ C. $\theta_w' = 0{\cdot}005^\circ$ C.

P.	C_n.	K_n.	C_w.	K_w.
77·56	0·22628	0 00005838	0·21013	0·00005947
64·66	0·22623	5836	0·20914	5892
54·43	0·22618	5833	0·20849	5856
45·26	0·22618	5833	0 20818	5838
36·19	0·22607	5828	0·20800	5828
32·00	0 22607	5828	0·20794	5825
27·94	0·22603	5825	0·20789	5822
23·73	0·22598	5822	0·20789	5822
19·24	0 22588	5817	0·20783	5818
15·21	0 22578	5812	0·20774	5814
10·71	0·22563	5805	0·20760	5805
5·61	0·22519	5781	0·20723	5785

$\theta_m = 10{\cdot}700^\circ$ C. $K_m = 0{\cdot}00005824$. $\alpha = 1{\cdot}1831$.

Table IV.—Nitrous Oxide.

$\theta = 17{\cdot}316^\circ$ C. $(R/l)_\theta = 0{\cdot}13334$. $\phi_\theta = 0{\cdot}00000535$ cal./cm.

$\theta_n' = 0{\cdot}008^\circ$ C. $\theta_w' = 0{\cdot}003^\circ$ C.

P.	C_n.	K_n.	C_w.	K_w.
81·54	0·16862	0 00003938	0·16316	0·00004358
62·09	0·16846	3931	0·16060	4222
52·10	0·16837	3926	0·15861	4117
43·05	0·16826	3921	0·15680	4023
37·24	0·16816	3917	0 15621	3992
34·31	0·16811	3914	0·15590	3976
31·46	0 16806	3912	0·15569	3966
28·49	0 16802	3910	0·15549	3956
25·63	0·16796	3907	0·15530	3946
22·76	0·16791	3905	0·15515	3938
19 81	0·16781	3900	0·15501	3931
16·87	0·16770	3896	0·15486	3923
13·65	0·16761	3891	0·15471	3916
10·58	0·16747	3884	0 15445	3903
7·59	0 16731	3877	0 15420	3890
4·39	0·16691	3858	0·15375	3867
1·97	0·16586	3811	0·15285	3821

$\theta_m = 8{\cdot}661^\circ$ C. $K_m = 0{\cdot}00003891$. $\alpha = 1{\cdot}1829$.

Table V.—Nitrous Oxide.

$\theta = 21 \cdot 390^\circ$ C. $(R/l)_\theta = 0 \cdot 13533$. $\phi_\theta = 0 \cdot 00000712$ cal./cm.
$\theta_s' = 0 \cdot 010^\circ$ C. $\theta_w' = 0 \cdot 004^\circ$ C.

P.	C_s.	K_s.	C_w.	K_w.
79·56	0·18662	0·00003962	0·18146	0·00004428
65·70	0·18653	3956	0·17916	4316
51·23	0·18634	3949	0·17656	4190
40·41	0·18618	3943	0·17362	4051
31·28	0·18602	3937	0 17237	3992
28·51	0·18598	3935	0·17211	3981
25·81	0·18593	3932	0·17191	3971
22 43	0 18588	3931	0·17167	3960
19·41	0·18578	3925	0·17147	3950
16·59	0·18573	3924	0·17127	3941
13 77	0·18552	3915	0·17111	3934
10·86	0·18542	3911	0·17085	3923
8 14	0·18523	3902	0·17062	3910
5 55	0·18493	3891	0·17031	3897
2·63	0·18403	3853	0·16951	3860

$\theta_m = 10 \cdot 699^\circ$ C. $K_m = 0 \cdot 00003924$. $a = 1 \cdot 1830$.

Table VI.—Nitrous Oxide.

$\theta = 25 \cdot 479^\circ$ C. $(R/l)_\theta = 0 \cdot 13731$. $\phi_\theta = 0 \cdot 00000890$ cal./cm.
$\theta_s' = 0 \cdot 012^\circ$ C. $\theta_w' = 0 \cdot 005^\circ$ C.

P.	C_s.	K_s.	C_w.	K_w.
75·48	0·20321	0·00004001	0·19783	0·00004482
64·50	0·20310	3997	0·19567	4384
53·60	0·20299	3993	0·19368	4295
42·05	0·20276	3983	0·19037	4147
32·80	0·20255	3975	0·18797	4043
29·65	0·20250	3973	0·18752	4024
26·97	0·20245	3970	0·18722	4010
24·62	0·20240	3968	0·18692	3997
22·61	0·20240	3968	0·18677	3991
20·46	0·20230	3964	0·18662	3985
18·38	0·20220	3961	0·18647	3978
16·60	0·20211	3957	0·18632	3972
14 22	0·20201	3953	0·18616	3966
12·40	0·20192	3949	0·18600	3959
10·75	0·20176	3943	0·18586	3952
8·82	0·20160	3936	0·18571	3946
6 73	0·20140	3929	0·18541	3933
4 70	0·20111	3917	0·18512	3920
2·65	0·20040	3889	0·18452	3894

$\theta_m = 12 \cdot 744^\circ$ C. $K_m = 0$ 00003960. $a = 1 \cdot 1831$.

The curves in fig. 1 refer to the variations of the apparent conductivity with pressure at constant temperature of carbon monoxide, and in fig. 2 of nitrous oxide. Figs. 3 and 4 indicate the manner in which the thermal conductivity of carbon monoxide and nitrous oxide respectively varies with temperature, the intercept of the line with the zero temperature axis, in each

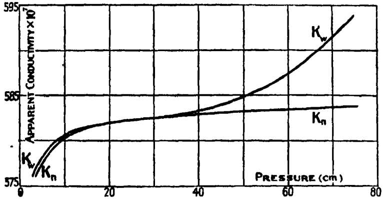


Fig. 1.—Carbon Monoxide.

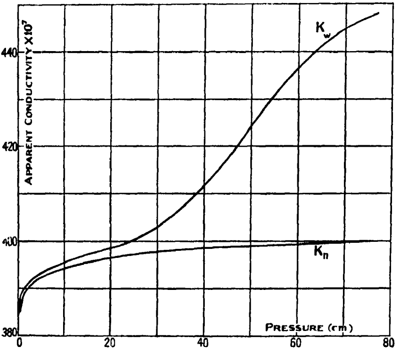


Fig. 2.—Nitrous Oxide.

case, giving the absolute value of the conductivity at 0° C. From the latter also are obtained the values of the coefficient of increase of conductivity, with temperature for each gas.

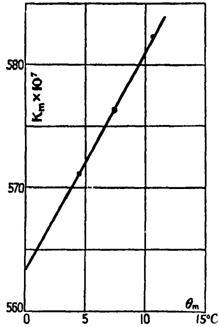


Fig. 3.—Carbon Monoxide.

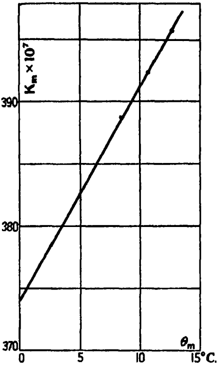


Fig 4.—Nitrous Oxide.

The results obtained were as follows :—

——	K_0.	Temperature coefficient.
Carbon monoxide	0·00005633	0·00317
Nitrous oxide	0·00003740	0·00462

In applying the results of the experimental determinations of thermal conductivities described above to the evaluation of the factor "f" in the relation

$$f = K/C_v\,\eta$$

where K, C_v and η are the thermal conductivity, the specific heat at constant volume and the coefficient of viscosity respectively, the difficulty of choosing appropriate values of C_v is apparent, owing to the relatively few direct experimental determinations of this quantity. The authors have endeavoured, however, to maintain consistency by calculating the values of C_v at 20° C.

from the results of experimental determinations of the specific heat at constant pressure (C_p), and of the ratio of the specific heats (γ) of the gases concerned in the neighbourhood of 20° C., for which reliable data are available. The thermal conductivities at this temperature have been taken from the authors' results, and the values of the viscosities from the results of Smith's experiments. The following table embodies the results :—

—	$K_{20} \cdot 10^7$.	C_p.	γ.	C_v.	$\eta_{20} \cdot 10^4$.	f.
N_2	613·9	0·2495	1·4045	0·1777	1·755	1·97
CO	599	0·2505	1·41	0·1776	1·755	1·92
CO_2	379·8	0·2020	1·303	0·1550	1·462	1·68
N_2O	408·6	0·210	1·280	0·1641	1·462	1·70

Thermal Conductivity of Air.

In view of the importance of the thermal conductivity of atmospheric air and as an additional check on the consistency of the observations detailed above with those undertaken by the authors in 1926, a series of measurements of the thermal conductivity of air was made, using the same apparatus as in the case of carbon monoxide and nitrous oxide, and using also another system of tubes constructed for the investigations of the thermal conductivities of vapours.

The results, which are in very close agreement with those obtained previously by the authors, are shown in the following table, along with data relating to the dimensions, etc., of the three systems of tubes. In this table the apparatus used in the work described above is indicated as Apparatus 1, the tubes constructed for the work on Vapours as Apparatus 2, and the original system used by the authors as Apparatus 3.

Dimensions, etc.	1.	2.	3.
	Cm.	Cm.	Cm.
Narrow tubes, internal radius (r_2)	0·5874	0·5831	0·4592
,, ,, external ,, (r_4)	0·6439	0·6346	0·5215
Wide ,, internal ,, (r_3) .	1·402	1·4033	1·1714
,, ,, external ,, (r_5)	1·456	1·4925	1·2360
Radius of platinum wire (r_1) .	0·0050596	0·0050655	0·007606
Constant $a = \dfrac{\log_e r_3/r_1}{\log_e r_2/r_1}$.	1·1830	1·1850	1·2289
Disposition of apparatus	Vertical.	Vertical	Horizontal.
Absolute conductivity at 0° C.	0·00005841	0·00005820	0·0000583
Mean value	0·0000583		

Studies on Fluorescence and Photosensitization in Aqueous Solution.
I.—Introduction.

By William West, Ph.D., Ralph Holcombe Müller, Ph.D., and Eric Jette, Ph.D., *Washington Square College, New York University, New York City.*

(Communicated by James Kendall, F.R S —Received July 11, 1928.)

It is well known that the concept of "collisions of the second kind," first enunciated in precise terms by Klein and Rosseland,* has been in the hands of Franck and his co-workers an instrument of great power in forming a theory of the possible ways in which an energy-rich atomic system may assume its normal state otherwise than by radiating its excess energy. This kind of collision is one in which an atom or molecule, which has by some means been raised from its normal to a higher quantum state, loses this additional energy in the course of a collision with a free electron or some other atom or molecule, which in its turn receives an increase in kinetic energy, or assumes a higher state of internal energy. The development of the idea has led to the prediction and confirmation of a number of phenomena in dilute gases, such as sensitized fluorescence and photochemical sensitization, which, taken together, form a firm experimental basis for some of the fundamental concepts of the quantum theory.

Now since it is universally believed that in a chemical reaction the true reactants are molecules which have in some way or other become "activated," and since there is at present almost complete unanimity that it is necessary to invoke the aid of collisions to account for the velocity of even a unimolecular reaction, it becomes relevant to examine the possibility of extending the concept of collisions of the second kind to the general problems of chemical reaction. Indeed, the essence of the idea appeared in the chemical literature long before its formal announcement and proof by Klein and Rosseland.†

---

* 'Z. Physik,' vol. 4, p. 46 (1921).

† See, for instance, Bodenstein, 'Z. Elektrochemie,' vol. 22, p. 58 (1916). Bodenstein attempted to account in the hydrogen-chlorine reaction for the large yield of hydrogen chloride per quantum absorbed by assuming a transfer of energy from the highly energetic "nascent" molecules of hydrogen chloride to molecules of reactants, a hypothesis which has appeared, in more quantitative form, under the guise of "reflex activation." See Christiansen and Kramers, 'Z. Phys. Chem.,' vol. 104, p. 457 (1923), Watson, 'Roy. Soc. Proc.,' A, vol. 108, p. 132 (1925).

It is, however, in the study of sensitised photo reactions that an examination of the utility of the concept in chemical questions can best be made, for in such a case we know with considerable precision, from the absorption spectrum of the sensitiser, the energy available for transference to the reactant by a collision of this kind. Several workers have already indicated in general terms the way in which the concept may be extended to these reactions.* In the present communications an attempt has been made to find evidence by means of quantitative experiments which would show whether the concept could be extended to solutions. There is no reason to suppose that atomic or molecular activation in liquids is fundamentally different from that in the gaseous phase, and we might expect that all the phenomena accounted for by collisions of the second kind in gases, namely, the extinguishing of fluorescence by added substances and by an excess of the fluorescent substance itself, sensitized fluorescence, and sensitized photochemical reactions, should find their counterparts in liquids. Effects due to collisions will, in fact, be at a maximum in the liquid state, and to these collisions must be attributed the great complexity of absorption and fluorescent spectra in liquids. A fluorescent band in a liquid will be a complete electronic- vibrational- rotational band, quite analogous to the iodine fluorescence spectrum in the presence of foreign gases, in which, however, all trace of fine structure has been lost owing to the smallness of the rotational quantum jumps associated with the large moments of inertia of the complicated fluorescent molecules, and to the displacements brought about as an inner Stark effect by the very large intermolecular fields prevailing in the neighbourhood of the molecules.

The most direct way, therefore, of obtaining information about these collisions in solution is to study the effect of adding various substances on the fluorescence of some substance in a liquid system. In this investigation certain salts have been found to exert an extinguishing effect on the fluorescence of several substances in aqueous solution. The extinguishing powers of the salts are in the same order for fluorescent substances of quite diverse chemical natures, and the order is found to be identical with the order of "deformability" of the ions as measured by such considerations as the Rydberg corrections for the spectral terms, or molecular refraction and dispersion. As will be shown in detail, the best explanation of the effects is to be found in collisions of the second kind, and these collisions are found to be the more probable, the more loosely bound the outer electrons of the added ion.

---

* Bodenstein, 'Trans. Faraday Soc.,' vol. 21, p. 531 (1926); Berthoud, *ibid.*, vol. 21, p. 554 (1926); J. Perrin, 'C. R.,' vol. 184, p. 1097 (1927).

We are still uncertain of the fate of the energy lost by the excited fluorescent molecule when the fluorescence is extinguished by salts; since, however, we find that the most deformable ions are the most effective, the first supposition is that the energy of the excited fluorescent molecule is used up in producing a deformation or polarisation in the ion, and experiments are in progress to test this conclusion.

Under certain conditions the energy may be utilized in the production of chemical changes; these may be regarded as cases of polarization so intense that, in the presence of a suitable acceptor, an electron is removed from the sensitized molecule on collision with the molecule primarily activated. We are then dealing with a case of photochemical sensitization, *i.e.*, the presence of the light-absorbing molecule causes, on illumination, a chemical change in a molecule insensitive in itself to the radiation in question. It is not to be supposed that only fluorescent substances will act as photosensitizers, substances capable of absorbing radiation of the appropriate frequency may be effective, whether fluorescent or not. There is indeed some reason for believing that, other conditions being equal, a non-fluorescing molecule might be a more efficient sensitizer than a fluorescent one. It would seem that any undisturbed atom or molecule would fluoresce after it had absorbed light, provided the absorption is not attended with disintegration of the molecule, and we must attribute the comparatively rare appearance of fluorescence in solution, despite the presence of an absorption band, to the great susceptibility of the excited molecules to radiationless transfers of energy on collision with solvent and solute molecules. Fluorescence in solution, then, seems to imply a certain stability of the activated molecule, a certain ability to resist collisions of the second kind. What evidence there is is favourable to this view. It has been found that the period for which fluorescence persists after cessation of the exciting light is the longer, the more intense the fluorescence.* Again, an observation of Winther's has been confirmed in this work, that Eder's reaction, the action between mercuric chloride and ammonium oxalate, is sensitized to visible light by the addition of small quantities of erythrosin and eosin, and not noticeably by fluorescein. Erythrosin is somewhat more effective than eosin. Of these compounds, fluorescein shows the most intense fluorescence and erythrosin the least. In the sensitization of the photographic plate by these closely related substances, the same order is observed—erythrosin is the most effective and fluorescein the least.

If then we adopt the hypothesis that photochemical sensitization is due essentially to the transfer of energy absorbed from the radiation by the sensitizer to the reacting molecule in a collision of the second kind, we can predict several effects. In the first place, if the sensitizer is fluorescent, we should expect its fluorescence to be extinguished in whole or part by the addition of the reacting substance. Also, if we add to the sensitized reaction mixture an ion which inhibits fluorescence by collision of the second kind, we have as competitors for the energy both the reactant and the additional ion. The amount of reactant changed in a given time of illumination should therefore be smaller in the presence of the ion, and there should be a parallelism between the extinguishing effect of different ions and their inhibiting action on the sensitization. Moreover, if the study of the extinguishing of fluorescence permits us to draw general conclusions concerning the factors which determine collisions of the second kind, we should expect the addition of ions to a system sensitized by a non-fluorescing substance to have the same kind of effect as when the sensitizer is fluorescent.

---

* Gaviola, 'Z. Physik,' vol. 42, p. 862 (1927).

We may say in short that all of these predictions have been fulfilled The photo-decomposition of aqueous oxalic acid solution has long been known to be sensitized for visible and long ultra-violet light by uranyl salts. It has been found that the addition of chloride, bromide, thiocyanate and iodide ions inhibit the sensitized photo-reaction, chloride being least effective. The order of the ions is the same as that in the inhibition of the fluorescence of uranyl ion. Uranyl ion has also been found to photosensitize the decomposition of glucose in solution. Glucose also extinguishes the fluorescence of uranyl nitrate, and the extent to which it does so is parallel to its capacity to undergo chemical change in the presence of light and uranyl ion. Tests have been carried out on Eder's reaction, sensitized by eosin. The mixture of ammonium oxalate and mercuric chloride inhibits the fluorescence of eosin, and at the same time undergoes chemical change on exposure to visible light. Sulphate, nitrate and fluoride, which have little effect on fluorescence, have little inhibiting power on the photosensitization, while chloride, bromide and thiocyanate, which inhibit the fluorescence to increasing extents in that order, affect the photosensitization in the same way The parallelism between the effect of ions on fluorescence and on photosensitization is exact. Finally, the effect of these ions on Eder's reaction sensitized by the non-fluorescent ferric ion has been found exactly to parallel the effect on fluorescent sensitizers. The evidence for regarding the extinguishing of fluorescence in solution by added ions and photochemical sensitization in solutions as effects of collisions of the second kind must therefore be regarded as strong.

The main outstanding problem in this field is the elucidation of the conditions which determine the probability of these collisions. In the case of the ions studied here the deformability of the outer electron shell has been shown to be a factor of first importance in determining this probability; but as yet there are practically no data which enable one to include non-ionic substances. The probability of these collisions depends upon both the excited molecule and the other colliding particle—a fact well enough illustrated in this work, where it has been found that although the extinguishing effect of ions is in the same order for different fluorescent substances, the absolute concentrations of the same ion required to produce a given extinction in different fluorescent compounds vary greatly from compound to compound. For photochemical sensitization a necessary condition is that the sensitizer should absorb a quantum sufficient to bring about the chemical change in the reactant, and, by analogy with the corresponding case in gaseous systems, the transfer of energy from the sensitizer will be the more probable the nearer the value of the energy of excitation to that required to bring about the chemical change in question. It is evident, however, that this is not a sufficient condition for transfer of energy, for experiment shows that a group of substances with absorption bands in approximately the same position and, therefore, with about the same energy of excitation are not equally efficient sensitizers for a given reaction. Thus fluorescein has little or no sensitizing effect in Eder's reaction, while the related compounds eosin and erythrosin, with absorption bands displaced somewhat toward the red, but not very different from that of fluorescein, are efficient sensitizers. Ferric ion, on the other hand, which absorbs higher frequencies than fluorescein, is a powerful sensitizer in this reaction. Quinine bisulphate, again, which absorbs in the long ultra-violet, has little or no sensitizing effect. Further experimental work, in which the sensitizing effects of different groups of chemically related compounds are examined, is required to throw more light on the conditions which determine these energy transfers; but even in the cases of comparatively simple gaseous fluorescence at low pressure the factors which determine the probability of collisions of the second kind are still little known.*

Summary.

The application of the concept of "collisions of the second kind" to processes occurring in solution, especially to extinction of fluorescence and photosensitization, is discussed, and the conditions necessary to be fulfilled for photochemical sensitization are analysed.

---

* Mankopff, 'Z. Physik,' vol. 36, p. 315 (1926).

Studies on Fluorescence and Photosensitization in Aqueous Solution.
II.—*Fluorescence in Aqueous Solution.*

By Eric Jette, Ph.D., and William West, Ph.D., Washington Square College, New York University, New York City.

(Communicated by James Kendall, F.R.S.—Received July 11, 1928.)

In the present work we have attempted to make a general systematic survey of the effect of the addition of salts to fluorescent solutions. The effect of the presence of salts on fluorescent solutions was noticed by the earliest workers on the subject; Stokes,* for instance, reports that while solutions of quinine in sulphuric, nitric, tartaric, succinic and phosphoric acid were fluorescent, the addition of the halogen hydracids and of sodium chloride extinguished the fluorescence. Similar observations were made by Buckingham† and recently by Desha, Sherrill and Harrison.‡

Apparatus.—We chose as measuring instrument the photoelectric cell. After preliminary experiments had shown that the inconstancy of the light source made it troublesome and inaccurate to measure the current from a single photo-cell illuminated directly by the fluorescent light, we finally adopted a differential method, the essentials of which are sketched below.

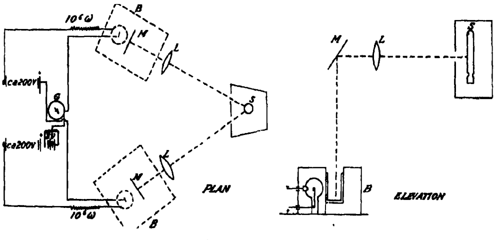


Fig. 1.—Apparatus.

---

* 'Phil. Trans.,' vol. 142, p. 541 (1852).

† 'Z. Phys. Chem.,' vol. 14, p. 129 (1894).

‡ 'J. Amer. Chem. Soc.,' vol. 48, p. 1493 (1926).

Two gas-filled potassium surface photo-cells, with characteristics as similar as possible, were connected in opposite ways through the same galvanometer. In front of each cell was a rectangular glass vessel containing the fluorescent solutions, illuminated by equal beams from an upright mercury vapour lamp in a manner made sufficiently obvious by the diagrams. One solution, the "standard," was kept unchanged during an experiment; the other, the "working solution," was varied in the desired way. The lenses (L) and the mirrors (M) were of glass, so that the shortest wave-length reaching the fluorescent solution would be about 3300 Å. The glass vessels, which were as nearly as possible identical, fitted snugly into blackened brass containers enclosing them on three sides and fashioned so as to be placed readily and accurately in position in front of the photo-cells through a slit in the lid of the blackened wooden box (B). The current and voltage of the lamp were read by the necessary instruments, as were also the potentials applied to the photo-cell. The whole of the optical train and the measuring instruments were rigidly mounted on a board so as to prevent displacement of the parts.

This arrangement, although not a null method, and though still showing the effects of fluctuations in the intensity of the source when there is a considerable difference in the intensities of fluorescence in the standard and working solutions, is a great improvement on the direct use of a single cell in that it is most precise in the region where greatest precision is required, *i.e.*, for small differences in intensity between the standard and the working solution. The simpler method, on the other hand, is subject to greatest error in this region. We assured ourselves by preliminary measurements that the photo-currents yielded by the cells were linear functions of the intensity of illumination within the range with which we were concerned.

The procedure was as follows : The mercury lamp having become constant, the standard and working vessels, containing 25 c.c. of pure fluorescent solution and 25 c.c. of distilled water, were placed in front of the photo-cells and the potentials on the latter adjusted until the galvanometer reading was zero. The working vessel was then withdrawn and filled with 25 c.c. of fluorescent solution and 25 c.c. of the salt solution whose effect was being examined. When the working cell was replaced, if the salt had any effect the photo-currents no longer compensated, and the galvanometer showed a deflection. This was repeated for various concentrations.

The effects of the salt are expressed as percentage extinctions of the fluorescence due to the pure fluorescent solution. To obtain the value for the pure fluorescent substance we placed in the standard vessel 25 c.c. of fluorescent material plus 25 c.c. of distilled water. In the working vessel was placed 25 c.c. of fluorescent solution to which was added 25 c.c. of salt solution of sufficient concentration to extinguish the fluorescence. Since the light reflected from the glass walls of the vessel and from the surface of the liquid would be the same in both cases, their effects would compensate, and, the presence of the salt not adding to the absorption in the wave-length region used in these experiments, the observed galvanometer deflection must then be a measure of the fluorescent light alone

The salts used were chemically pure preparations which were submitted to at least one crystallization. The fluorescent substances were also purified, where possible by recrystallization, but in some cases the contamination of the commercial samples by inorganic salts was so great that only by preparing and isolating the substances by methods excluding the use of inorganic salts was a satisfactory product obtained.

As the work of Desha, Sherrill and Harrison (*loc. cit.*) has shown the importance of the hydrogen ion concentration in determining the fluorescence of certain solutions, we exercised due precautions to prevent changes in this factor. Stock solutions of the fluorescent materials were kept in waxed bottles protected against carbon dioxide from the air, and the salt solutions were also kept in waxed bottles.

Results.

Quinine Bisulphate.—A 0·0025 molar aqueous solution of quinine bisulphate was used. The results are shown in Table I and fig. 2, in which the percentage extinction of the fluorescence is plotted against the concentration of the salt present in the solution; the concentrations are expressed in milli-equivalents per litre. In this series the positive ion, the potassium ion, was constant and the negative ion varied. It will be observed that the effect varies enormously with the salt; iodide, thiocyanate, bromide and chloride have a very great power of diminishing the fluorescence; sulphate, nitrate and fluoride have scarcely any effect, and oxalate and acetate occupy intermediate positions

We now performed a series of experiments in order to determine whether any difference could be detected in the behaviour of various positive ions. The effects of the chlorides of Li, Na, K, Rb and La were examined. The result was rather surprising; within the limits of error, equivalent concentrations of three different salts produced identical extinctions in the fluorescence of quinine bisulphate solution. (*Cf.* fig. 2.)

According to all modern conceptions of the constitution of salt solutions,

Table I.—Extinction of the Fluorescence of Quinine Acid Sulphate by the Addition of Salts.

Salt concentration in milli-equivalents per litre.	Percentage extinction										
	KNO_3	K_2SO_4	KAc	$K_2C_2O_4$.	KCl.*	KBr.	KCNS.	KI.	$HgCl_2$	$AgNO_3$.	AgAc
2·5	—	0·7	7·5	10·5	23·5	27·5	30·0	35·0	<0·5	4·5	8·0
5·0	—	1·2	14·0	19·5	38·0	42·5	45·0	53·0	0·8	6·5	15·5
10	—	2·2	24·5	35·0	55·5	61·5	63·5	72·0	1·5	8·5	22·0
15	—	3·1	33·5	46·5	66·0	71·0	72·5	81·0	2·2	10·0	36·5
25	0·5	4·8	45·0	59·5	76·5	81·5	83·5	89·5	3·5	12·0	51·0
50	1·0	8·5	56·5	75·0	84·0	90·5	91·5	96·0	5·7	13·5	72·0
100	—	10·8	62·0	81·0	92·5	95·0	96·0	100	10·1	22·5	—
500	<2·0	17·0	65·0	—	97·5	99·0	100	—	25·0†	—	—
1000	—	20·2	65·0	—	99·0	100	—	—	—	—	—

* Same curve for $LaCl_3$, LiCl, NaCl, RbCl

† Concentration 400 milli-equivalents

we must attribute the extinguishing of the fluorescence to the ions from the added salt. Since it is much more likely that the positive ions mentioned above are almost entirely without effect on the fluorescence than that particles which differ so much in size and charge should have identical effects, we are led to

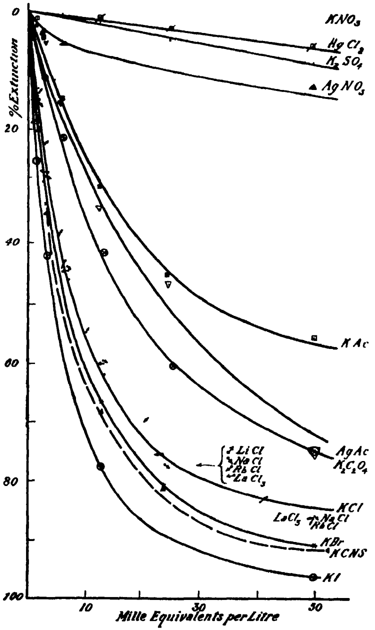


Fig. 2.—Effect of Salts on the Fluorescence of Quinine Bisulphate.

conclude that only the negative ion is effective. As will be mentioned later, other cases, of at first sight quite a different nature from this studied here, are known in which negative ions exert large influences and positive ions are nearly without effect.

In order to verify if the presence of ions was essential for the extinguishing of the fluorescence, we observed the effect of the addition of mercuric chloride, a " salt " which is known to be very slightly ionized in solution at concentrations at which the chlorides previously used are practically completely ionized. The curve is shown in fig. 2, from which it is evident that, under comparable conditions, the effect of mercuric chloride is much smaller than that of the other chlorides. The apparent degree of electrolytic dissociation of $HgCl_2$ at 0·1 normal concentration is less than 1 per cent. A solution of KCl at 0·1 normal concentration diminished the fluorescence intensity by 92·5 per while the same concentration of $HgCl_2$ diminished it only by approximately 10 per cent.* Since the concentration of Cl^- ions in the latter solution was less than 0·001 normal, inspection of the curve shows that this was approximately the diminution to be expected, if the effect is due essentially to the chloride ions. Mercuric nitrate and acetate each react with quinine bisulphate with the formation of a yellow precipitate, hence it was impossible to observe the effect of mercuric ion from a highly ionized mercuric salt. The ions $Li^\cdot$, $Na^\cdot$, $K^\cdot$, $Rb^\cdot$, $La^{\cdots}$, each possess the rare gas structure of the outer electronic shell. The silver ion, however, has an outer shell of 18 electrons, and in order to find if this type of electronic configuration has a different effect on fluorescence, the light emitted by the quinine bisulphate in the presence of silver nitrate was measured. This substance was found to cause a distinctly greater diminution of fluorescence than potassium nitrate (see fig. 2). The effect of silver ion was confirmed on experiment with silver acetate, the curve for which was found to fall below that of potassium acetate.

Zinc and cadmium ions (from the nitrates) while possessing the same outer electronic configuration as silver ion, have no measurable effect on the fluorescence, which shows that this particular configuration is not in itself sufficient to account for the effect; magnesium nitrate also had no extinguishing effect.

A comparison of the curves for potassium acetate, silver nitrate and silver acetate suggests that the ionic effects may be approximately additive in those cases where both ions of a salt extinguish the fluorescence.

---

[*] The difference between the effects of $HgCl_2$ and KCl on the fluorescence of quinine bisulphate has also been noted by Buckingham.

Finally, as examples of non-electrolytes, solutions of cane sugar and of urea were added. The effect of both substances was very slight; a 0·95 normal solution of cane sugar diminished the fluorescence 2·86 per cent. and a 1·89 normal solution by 5·72 per cent. It was found, moreover, in a qualitative experiment that the inhibiting effect of potassium bromide on the fluorescence of quinine bisulphate was distinctly less when the solvent for the substances was a very concentrated, viscous solution of cane sugar, than when water was the solvent. These effects may perhaps be accounted for by the high viscosity of these solutions.

Uranine.—The substance was prepared in a pure condition by hydrolysis of purified fluorescein diacetate,* which was in turn prepared from a good commercial sample of fluorescein. A solution of the disodium salt was used in the experiments, obtained by dissolving the acid in the necessary amount of sodium hydroxide solution. The concentration of the solution used was 0·00025 molar. The effects on this substance were much the same as on quinine bisulphate, but much higher concentrations were required to produce the same percentage of extinction. Solutions of potassium nitrate, sulphate, oxalate and chloride at concentrations as high as 1·00 normal had no measurable effect. The silver ion, again, had an unusually large effect in concentrations less than N/10, above which silver ion causes precipitation of the fluorescein. The results are given in Table II and fig. 3; the units used are the same as in the preceding table and figure.

Table II.—Extinction of the Fluorescence of Disodium Fluorescein by the Addition of Salts.

Salt concentration in milli-equivalents per litre.	Percentage extinction.			
	KBr.	KCNS.	KI.	$AgNO_3$.
25	1·5	22·0	29·0	50·0
50	3·5	38·0	44·0	62·5
100	6·0	57·5	65·0	—
250	12·5	80·0	83·0	—
500	19·0	90·5	94·5	–
1000	27·0	93·0	98·9	—

---

* Orndorff and Hemmer, 'J. Amer. Chem. Soc.,' vol. 49, p. 1272 (1927).

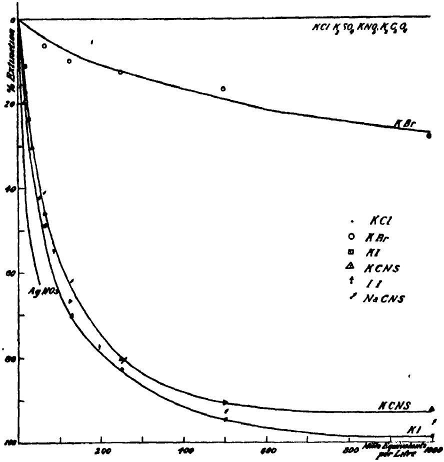


FIG. 3.—Effect of Addition of Salts on Fluroescence of Disodium Fluorescein.

Uranyl Salts.—The uranyl salts are particularly interesting owing to their known photo-catalytic properties. A survey of some common orange salts showed that the sulphate and nitrate fluoresced quite brightly, with a green colour, the acetate less, and, significant from our point of view, the chloride very feebly. But the intensity of fluorescence even of the sulphate, which was chosen for investigation, is considerably feebler than that of uranine and is in a wave-length region to which the potassium surface photo-cell is not very sensitive. Accordingly, we made no effort in this case to determine the course of the extinction curve, but contented ourselves with ascertaining the order of the effect of various salts by noting the volume of salt solution which had to be added in order to diminish the fluorescence nearly to zero. A cell containing N/10 uranyl sulphate was placed in a box and illuminated by an intense beam of light from a carbon arc, which was passed through a condensing lens, a water cell, and a blue glass filter. Salt solution was added from a burette until the fluorescence, observed perpendicularly to the exciting beam, practically disappeared. The results are collected below :—

Salt added.	*Volume required to reduce fluorescence to zero.*
Sodium fluoride N	Addition of a few cubic centimetres markedly increased the fluorescence. Further addition weakened it somewhat from the increased value, but the fluorescence was still bright after addition of 25 c.c
Potassium sulphate N/1	Addition of 25 c.c. had little effect.
Potassium nitrate N/10	,, ,,
Potassium acetate N/10	,, ,,
Potassium oxalate N/10	28 c.c. reduced fluorescence to zero.
Potassium chloride N/10	11 c.c. ,, ,,
Potassium bromide N/10	1 c.c. ,, ,,
Potassium thiocyanate N/10 ..	0·5 c.c. ,, ,,
Potassium iodide N/10	0·30 c.c. ,, ,,
Mercuric chloride	Addition of 30 c c. 0·4 normal $HgCl_2$ diminished the fluorescence only by about 2/3.

In all these fluorescent substances there is a well-defined order in which negative ions exert an inhibiting effect on the fluorescence. Beginning with the most powerful inhibitor, the series is

$$I' > CNS' > Br' > Cl' > Ox'' > Ac' > SO_4'' > NO_3' > F'.$$

There are, however, a number of fluorescent substances which are quite unaffected by some of the most powerful inhibitors of the above series, and largely influenced by the nitrate ion which has little effect on the three substances mentioned above. On investigating the effect of salts on the fluorescence of the sodium salts of beta-naphthol and 1-4 naphthalene disulphonic acid and on anthranilic acid, we found that potassium chloride, bromide and thiocyanate were without effect even to concentrations as high as 4 normal, iodide at high concentrations had in all cases a well-marked inhibiting effect, while potassium nitrate, one of the most inert ions for the former substances, was the most powerful inhibitant for these. In these compounds it is practically certain that the fluorescence originates in the aromatic nucleus. The fluorescence of benzene vapour has been found to be little influenced by the addition of foreign gases,* and evidently the stability of the fluorescent centre indicated by this fact persists in solution. The effect of the nitrate ion, however, is still unexplained.

Discussion.—In seeking for an explanation of the inhibiting effect of certain ions on fluorescence, we find that the most obvious possibility is that the added salt prevents the *excitation* of the fluorescent molecule to a condition in which it can re-emit its energy as radiation. This could happen in two ways: (1) If the salt possessed an absorption band in the neighbourhood of that of the fluorescent substance, it could act as an internal screen and prevent the excitation of the fluorescent molecules; or (2) some kind of "compound" could be formed between the added salt and the fluorescent substance which has not the properties, whatever they may be, which permit an energy-rich molecule to regain its normal state by the emission of radiation. We examined the absorption spectra of the salts at concentrations up to the greatest used in the fluorescence experiments and found that in no case did the absorption band extend into the region of the spectrum to which the fluorescent substance was exposed. The first possibility therefore falls to the ground. Information on the second was sought by examining the absorption spectrum of the fluorescent substance in the presence of the salt. For quinine, the small Hilger quartz spectograph E 31 was used, while for the visibly coloured substances the Hilger constant deviation glass spectograph was used. No material change in the absorption was observed, whence we conclude that the addition of the salt has produced no fundamental change in the constitution of the molecule of the fluorescent substance, and, further, that this molecule is still capable of undergoing, in the presence of the added salt, the transition from the normal to the excited state, the return from which, in the absence of the salt, is accompanied by the emission of radiation. The return to the normal condition must then take place through a radiationless transfer of energy to some other molecule in the system in the course of a "collision of the second kind."

The diminished inhibiting power of ions in very viscous solutions is in complete accord with the interpretation in terms of collisions. Frances Perrin† has shown that with increased viscosity of the solvent, the concentration of fluorescent substance at which maximum fluorescence occurs will increase.

---

* Pringsheim, in Geiger and Scheel's "Handbuch der Physik," vol. 23, "Quanten," chap. 5, p. 529.

† 'C. R.,' vol. 178, p. 2252 (1924).

Now we must ascribe the diminishing fluorescent power of a substance with increasing concentration to collisions of the second kind between the fluorescent molecules themselves, and evidently, with an increase in viscosity of the medium, the frequency of collision between fluorescent molecules is diminished. The same is true for collisions with other molecules. The effect of increased viscosity in diminishing the inhibiting powers of ions is thus of exactly the same nature as its effect in increasing the concentration at which maximum fluorescence of the pure fluorescent substance occurs.

For the three fluorescent substances under discussion the fluorescence must be ascribed mainly to fluorescent ions in solution. Quinine bisulphate, the sodium salt of fluorescein, and uranyl sulphate are typical salts and as such must be nearly completely ionized at the concentrations employed. It was therefore possible that the effect of the added salt might be immediately explicable in terms of the modern thermodynamic theory of ionic solutions, and, in particular, as an effect of the salt on the thermodynamic activity of the fluorescent ion. The absence, however, of differences in the effects of positive ions (except in the case of silver ion) is strongly against there being any direct effect of this kind; the trivalent lanthanum ion, for instance, ought to have had a markedly different effect from the monovalent rubidium ion, which is contrary to experience. In short, the evidence is against there being any direct relation between the different effects of ions on fluorescence and the following concepts in the theory of solution: (1) Differences in ionic size, whether in the hydrated condition or not; (2) difference in the activity coefficients of the individual ions of the salt added; (3) differences in the changes in the activity coefficients of the fluorescent ions produced by the added salts; (4) differences produced by the added salts in the degree of hydrolysis of the fluorescent salts.

The order in which ions suppress the fluorescence of the substances discussed in this report is very similar to the well-known lyotropic series of ions which appear in many diverse fields of physical chemistry. These salt series further resemble that which we have found to hold for fluorescence in that the differences between positive ions are frequently much less marked than those between negative ions.

The parallelism between the lyotropic series as usually found and that observed in fluorescence is, however, not exact. Thiocyanate in the lyotropic series usually comes beyond iodide, while we have consistently found its effect on fluorescence to be between that of bromide and iodide; and nitrate, which in the lyotropic series occurs between bromide and iodide, is in ours at the opposite end of the series from iodide. As we shall indicate, we believe these differences to be significant.

Heydweiler* in 1913 pointed out that the order of the optical dispersions of negative ions for visible light was parallel to that of their precipitating power in colloid chemistry, *except that thiocyanate, instead of being beyond iodide, was intermediate between iodide and bromide.* The molecular refraction is expressed by the Lorenz-Lorentz equation

$$R = (n^2 - 1)/(n^2 + 2) \, . \, M/\rho$$

in which n is a value of the refractive index for infinitely long wave-length found by extrapolation of the dispersion curve in the visible. R is a measure, in the first place, of the polarization produced in the atoms by the displacement of the electrons under the influence of the electro-magnetic field, and hence, in general, of the ease with which the electrons are displaced by electrical forces. The greater the refraction the greater the deformation produced in an atom or ion by the displacement of its outer electrons, or the looser the binding of the electrons. In recent years the work, particularly of Fajans and Joos, of Smyth, of Born and Heisenberg† and others, has emphasized the importance of the information to be obtained about the forces concerned in atomic and molecular structure from the study of atomic and ionic deformabilities. The order of the deformability of negative ions on the basis of these considerations is

$$F' < NO_3' < SO_4'' < CO_3'' < Cl' < I' < CN' < O'' < S'' < Se''‡$$

In the following table is arranged the order in which the ions extinguish fluorescence; under each ion is a number which is a measure of its deformability. For simple ions this is the refraction for the D line; for the nitrate and sulphate, the ionic refraction per oxygen atom has been used, the electrons of the central N and S atoms respectively being so tightly bound as to contribute practically nothing to the refraction for visible light (Fajans and Joos, *loc. cit.*). The interatomic bindings in the acetate, oxalate and thiocyanate are too uncertain to allow an average value of the refraction per atom to have much significance as a measure of the binding of the electrons; but the position of thiocyanate between bromide and iodide is indisputably given by the dispersion constants of these ions :—

---

* 'Ann. Physik,' vol. 41, p. 499 (1913).

† Fajans, 'Naturwiss,' vol. 2, p. 165 (1923); Fajans and Joos, 'Zeit. f. Physik,' vol. 23, p. 1 (1924); Smyth, 'Phil. Mag.,' vol. 50. p. 361 (1924); Born and Heisenberg, 'Z. Physik,' vol. 23, p. 388 (1924).

‡ Fajans, 'Naturwiss,' vol. 2, p. 165 (1923).

F'	NO_3'	SO_4''	$\overline{Ac''}$	$\overline{Ox''}$	Cl'	Br'	CNS'	I'
2·5	3·66	3·65			8·7	12·2		18·5

Dispersion constant $= (\Delta n_\gamma - \Delta n_\alpha)10^3$ for Br′, 73 ; CNS′, 100 ; I′, 175.

Δn_γ is a measure of the difference in refractive index between solution and solvent for the hydrogen line H_γ, and Δn_α this quantity for H_α. The value of the ionic refractions for some positive ions is also given for comparison.

$Li^{\cdot}$	$Na^{\cdot}$	$K^{\cdot}$	$Rb^{\cdot}$	$La^{\cdot\cdot\cdot}$	$Zn^{\cdot\cdot}$	$Ag^{\cdot}$
0·2	0·5	2·23	3·6	4	0·29	4·79

The high deformability of silver ion in comparison with the others is evident. It may also be mentioned that the absence of an effect of the positive ion in fluorescence is paralleled by a similar lack of contribution of positive ions to the dispersion in visible light ; the dispersion for visible light of different salts with a common anion is with some exceptions practically independent of the cation.*

A comparision of the deformability series and of the fluorescence series shows a parallelism too close to be accidental. We can state with certainty that the ions which exert the greatest inhibiting effect on fluorescence are those which are most deformable, *i.e.*, those in which the binding of the outer electrons is loosest.

It will be noticed that nitrate and thiocyanate ions are displaced in the Hofmeister and related series from the positions which they occupy in the deformability and fluorescence series. It is generally agreed that the action of the ions in the salting-out effect and in the related phenomena most probably takes place through their effect on the water dipoles in the solution. Now the nitrate and thiocyanate ions are likely to possess, in addition to the electric moment induced in the individual atoms by displacement of the electrons, a permanent dipole moment due to unsymmetrical distribution of charge over the ion as a whole. This would not contribute to the dispersion in the visible and hence would not appear in the deformability series which refers to the induced moment produced by electronic as distinct from atomic displacements ; but might be an important factor in determining the effect of an ion on other permanent dipoles like water molecules. The identity of the ionic deformability series and the fluorescent series must be taken to indicate a type of interaction between the extinguishing ion and the fluorescent substance in which the outer electrons of the former are primarily concerned; any nuclear vibration in the extinguishing ion is not affected by the fluorescent substance except in so far as it may be coupled with an electronic charge.

---

* Heydweiler and Grube, ' Ann. Physik,' vol. 49, p. 653 (1916).

The fate of the energy transferred from the fluorescent molecule to the extinguishing particle in a collision of the second kind is a question still unanswered. Since, however, we find that the most deformable ions are most effective in extinguishing fluorescence, our first hypothesis would naturally be that the energy goes to produce a deformation in the ion, or, more precisely, that an outer electron is displaced from its normal position to a position in which the force binding it to the core is weakened. The chemical activity of the ion will therefore be increased, and we can see how fluorescent substances might photo-catalyse actions depending on the presence of an "acceptor" for the electron (*i.e.*, oxidation-reduction reactions); for instance, how quinine illuminated by light might catalyze the oxidation of the iodide ion.[*] If the ions are really in some condition of strain such as we have imagined, their optical properties might be influenced by the addition of the fluorescent substance, and we have accordingly undertaken an examination of the refractive indices of mixture of salts and fluorescent substances in order to test the hypothesis.

Summary.

The effects of various salts on the fluorescence of aqueous solutions of three representative fluorescent substances, quinine bisulphate, sodium salt of fluorescein and uranyl sulphate have been quantitatively examined.

It has been found that :—

1. The extinguishing effect is associated practically entirely with the anion, silver ion being the only positive ion whose effect was noticeable.
2. The order of the extinguishing effect of the anions is

$$\mathrm{I}' > \mathrm{CNS}' > \mathrm{Br}' > \mathrm{Cl}' > \mathrm{Ox}'' > \mathrm{Ac}' > \mathrm{SO_4}'' > \mathrm{NO_3}' > \mathrm{F}'.$$

This series is identical with the deformability series of the ions.

The effects are interpreted in terms of the conception of collisions of the second kind, and the main result of the investigation may be summarised in the statement that the more deformable the electron orbits of the colliding particles, the more probable is a collision in which energy is transferred to it from particles in a higher quantum state than the normal.

---

[*] Pinnow, 'Ber. D. Chem. Ges.,' vol. 34, p. 2528 (1901).

Studies on Fluorescence and Photosensitization in Aqueous Solution.

III.—*Photosensitization and Fluorescence.*

Ralph Holcombe Müller, Ph.D., Washington Square College, New York University, New York City.

(Communicated by James Kendall, F.R.S.—Received July 11, 1928.)

It has been known for a long time that many photochemical reactions can be sensitized to a desired spectral region by the addition of a substance possessing an absorption band in that region. It is necessary that the quantum absorbed by the sensitizer is sufficient to bring about the chemical change in the reactant. Despite the wealth of experimental material and numerous theories offered in explanation, this class of reactions may be said to have found its first unequivocal interpretation in the classical experiment of Cario and Franck* on the dissociation of hydrogen by optically excited mercury atoms. The complete study of the problem showed that a mercury atom in the $2\,^3p_1$ state can either re-radiate its energy in the form of resonance radiation λ 2537 Å.U., or by collision yield the energy to a hydrogen molecule and dissociate the latter into atoms. One consequence of this picture is that for a given concentration of mercury vapour and constant intensity of the radiation 2537 Å.U., the yield of atomic hydrogen should increase with increasing hydrogen pressure. Another consequence is that the amount of re-radiated resonance radiation should decrease with increased hydrogen pressure.

Thus in the two reactions :—

$$(1)\qquad \dot{\mathrm{Hg}} \rightarrow \mathrm{Hg} + h\nu_{2537}$$
$$(2)\quad \dot{\mathrm{Hg}} + \mathrm{H}_2 \dashrightarrow \mathrm{Hg} + 2\mathrm{H}$$

the probability of reaction (1) taking place would be decreased and the probability of reaction (2) taking place would be increased by increased hydrogen pressure. This was found to be the case. Franck has pointed out that this situation is probably quite general, and offers a similar explanation for the sensitization of the photographic plate by means of dyes.

It has been suggested that sensitized photochemical reactions occurring in solution might be treated from the same viewpoint. There are several sensitizers which afford means for a direct test of this viewpoint. These are the sensitizers which exhibit fluorescence in solution.

---

* 'Z. f. Physik,' vol. II, p. 162 (1922).

Uranyl ions photocatalyse a number of reactions, among which the photolysis of oxalic acid is perhaps best known. The quantum efficiency for the photolysis of oxalic acid in the presence of uranyl ion has been determined by Büchi.[*] Starting with a definite amount of uranyl ion, the number of molecules of oxalic acid decomposed per quantum absorbed increases with increasing concentration of oxalic acid, approaching unity at a concentration of oxalic acid equal to that of the uranyl ion. According to Büchi the photolysis is to be attributed to the decomposition of the unionized molecules of uranyl oxalate, or of the complex ions $\mathrm{UO_2(C_2O_4)_2^{--}}$.

From our point of view, however, the absorption of a quantum of radiation by the uranyl ion leads to the formation of an energy-rich or excited ion,

$$\mathrm{U\dot{O}_2^{\cdot}} + h\nu \longrightarrow \mathrm{\dot{U}\dot{O}_2^{\cdot}}$$

which upon collision with oxalic acid or the oxalate ion transfers its energy to the latter, decomposing it and itself returning to the normal state. After a collision of this type the reverse of the above reaction, *i.e.*,

$$\mathrm{\dot{U}\dot{O}_2^{\cdot}} \longrightarrow \mathrm{U\dot{O}_2^{\cdot}} + h\nu$$

cannot take place because the energy of the excited ion has been lost to the oxalic acid. If this analogy with the case of excited mercury and hydrogen is warranted, then for a given concentration of $\mathrm{U\dot{O}_2^{\cdot}}$ the photochemical yield should increase and the fluorescence should decrease with increasing concentration of oxalic acid. In addition, the yield should approach a maximum, indicated by the complete disappearance of fluorescence, and then remain constant for concentrations of oxalic acid greater than this value. This was found to be the case.

Experimental.

The fluorescence of tenth molar uranyl sulphate solution was measured and compared with the fluorescence of the same solution containing definite amounts of oxalic acid. Mixtures were made up to a maximum ratio $\mathrm{H_2C_2O_4/U\dot{O}_2^{\cdot}} = 2$, the concentration of $\mathrm{U\dot{O}_2^{\cdot}}$ being constant throughout. Owing to the relatively feeble fluorescence of these solutions, the method described in the preceding paper could not be used. The photo-electric current was therefore amplified by means of a three-stage thermionic amplifier.[†]

The extinction of fluorescence by the addition of oxalic acid, determined in this manner, is shown in Table I and fig. 1. The dependence of the quantum

---

[*] 'Z. Phys. Chem.,' vol. III, p. 269 (1924).

[†] A new method for amplifying extremely small photo-electric currents has been developed by Mr. H. M. Partridge and the author and will be described shortly elsewhere.

Table I.

Ratio Oxalate/UO_2^{++}.	Fluorescence in Arbitrary units.*	Ratio Oxalate/UO_2^{++}	Fluorescence in Arbitrary units.*
0·10	1·15	0·80	0·27
0·25	0·85	1·00	Too feeble for accurate measurement.
0·50	0·54	2·00	,, ,,

* These values represent the values of the amplified photoelectric current expressed in milliamperes.

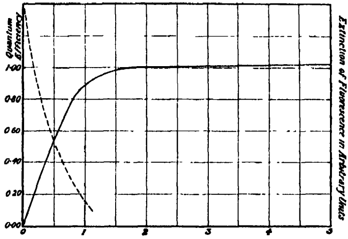


Fig. 1. -Average Concentration of Oxalic Acid.

efficiency upon the concentration of oxalic acid, as determined by Büchi, is included for comparison. The results are in accord with the hypothesis advanced above. The energy absorbed by a uranyl ion can be re-emitted as fluorescent light, or can be used to decompose the oxalic acid by a collision of the second kind. With higher concentrations of oxalic acid and, therefore, a higher collision probability, the probability for a photochemical change should be greater than for the emission of fluorescent light. This has been found to be true for most of the reactions sensitized by uranyl salts, a number of which are being studied quantitatively in this laboratory. Preliminary results on the photochemical decomposition of glucose in the presence of uranyl ion show that the yield of decomposition products runs strictly parallel with the extinction of fluorescence.

The crucial test for our application of the idea of "collisions of the second kind" to this problem should be to study the effect of adding ions, which are known to extinguish the fluorescence of uranyl salts, to the reaction mixture. The reactant and the added ions must then compete for the energy of the excited uranyl ion. Before giving the results let us note the order for the ions—

$$\bar{\mathrm{I}} > \overline{\mathrm{CNS}} > \bar{\mathrm{Br}} > \bar{\mathrm{Cl}} > \overline{\mathrm{C_2O_4}} > \overline{\mathrm{SO_4}}$$

according to their ability to extinguish fluorescence. The same order should hold for their power to inhibit the decomposition of the oxalic acid. It is interesting to note that Büchi found the reaction to be independent of added substances, but it so happened that the substances he chose have little effect upon the fluorescence of uranyl salts, and from our point of view this would lead one to expect very little inhibition of the photochemical reaction.

Fig. 2 and Table II show the effect of increasing amounts of "extinguishing ions" upon the photolysis of oxalic acid in the presence of uranyl sulphate. The concentrations employed were $\mathrm{U\ddot{O}_2}$ 0·01 molar; $\mathrm{H_2C_2O_4}$ 0·1 molar; the added ions ranging from zero to 0·2 molar. The reactions have been

Table II.

Molar concentration of added ion.	Decomposition of oxalic acid in per cent.			
	Cl⁻.	Br⁻.	CNS⁻.	I⁻.
0·02	27·1	26·2	23·0	14·0
0·05	26·1	24·2	18·5	9·1
0·10	24·8	22·3	13·8	6·2
0·20	24·2	19·4	8·6	4·0

Without added ion = 28·0 (av. of 4 det.).

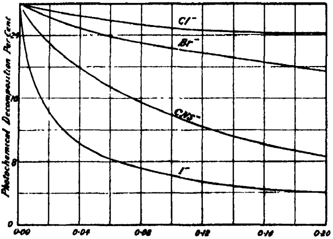


Fig. 2.—Concentration of Added Ion.

carried out at various times with light from a powerful incandescent lamp, a quartz mercury-vapour lamp, and sunlight In every case glass vessels were employed so that only the wave-lengths absorbed by the uranyl ion could be effective.

Chloride ion has a small inhibiting effect, bromide ion somewhat greater, thiocyanate ion a still greater and iodide ion the greatest effect The latter has been included only for its general interest in the series, for the results with this ion are not strictly comparable with the others because a small amount of free iodine was always liberated during the reaction. Correction for this fact still gives us an order for the inhibiting effect identical with the order for the extinguishing power.

The study of these inhibitor effects, together with the additional information furnished by the fluorescing sensitizer, gives us a powerful method of seeking to determine the fate of the absorbed energy We have found the same order to exist for the inhibition of reactions photo-catalyzed by non-fluorescing ions, notably in the case of Eder's reaction sensitized with ferric ion.

Further studies are under way to determine to what extent the energy escaping as fluorescent light can account for deviations from unit quantum efficiency, and also to find out how successfully the hypothesis can be applied to those sensitizers which give no "tell-tale" fluorescence but seem to dissipate the energy in some other way.

Summary.

1. The quantum efficiency for the photolysis of oxalic acid in the presence of uranyl ion approaches unity at a ratio $H_2C_2O_4/UO_2$ of 1 . 1, and remains constant for higher ratios. The fluorescence of uranyl ion with increasing additions of oxalic acid decreases, reaching zero at the same mole ratio. This is taken as good evidence for the similarity of this reaction to the case of excited mercury and molecular hydrogen studied in the gaseous phase by Cario and Franck.

2. The rate of decomposition is inhibited by those ions which possess the ability to extinguish the fluorescence of uranyl ion.

3. The order for the inhibiting effect and the extinguishing effect is the same for these ions.

4. The same order for the inhibition of photochemical reactions sensitized by non-fluorescing ions holds for these ions.

The Spectrum of Doubly-Ionised Nitrogen (N III).

By L. J. Freeman, B.Sc., D.I.C., Imperial College of Science and Technology, South Kensington.

(Communicated by Prof. A. Fowler, F.R.S.—Received July 27, 1928.)

Introductory.

Many of the lines which can now be attributed to doubly-ionised nitrogen (N III) have been known for a considerable time. The lines at wave-lengths 4097, 4103 and 4379 were recognised first by Lockyer, Baxandall and Butler* as being "enhanced" relative to the more familiar lines in the nitrogen (N II) spectrum. Careful measurements of these and other nitrogen lines were given by Clark† in 1914. In two papers by Fowler,‡ the wave-lengths of 23 lines were given which were observed to be enhanced in strong discharges in a vacuum tube containing nitrogen.

The first attempt to classify any of the lines was made by Fowler,§ who suggested that the lines λ 4103 and λ 4097 formed a principal doublet, and the three lines λ 4642, λ 4640 and λ 4634, a diffuse doublet. Following these suggestions, and utilising his own observations in the extreme ultra-violet, Bowen|| has classified 21 lines of the doublet system of N III and has given the values of eight terms, including the deepest of each of the three families of terms indicated by the Hund theory. The present paper may be considered as an extension of the identification of the predicted terms, including most of the strong quartet terms.

Observational Data.

The spectrum has been photographed from λ 8000 to λ 850 by using condensed discharges in vacuum tubes containing nitrogen. The lines of N III have been distinguished from those of other stages of ionisation by the usual method of comparing the relative intensities of the lines under various conditions of discharge. This method has to be applied with caution, for several of the more easily excited lines of N III (notably $2p\,{}^2P-2p'\,{}^2D$ at λ 990 and $2p'\,{}^2P-2p^3\,{}^2D$ at λ 1751) behave almost like N II lines.

---

* 'Roy. Soc. Proc.,' A, vol. 82, p. 532 (1909).

† 'Astrophys. J.,' vol. 40, p. 332 (1914).

‡ 'M. N. R. A.S.,' vol. 80, p. 692 (1920), and vol. 81, p. 189 (1921).

§ 'Report on Series in Line Spectra,' p. 165 (1922).

|| 'Phys. Rev.,' vol. 29, p. 231 (1927).

Below is a list of the instruments and dispersions available for investigating the various regions of the spectrum.

λ 8000—λ 5000	Glass spectrograph	25 A. per mm.	(mean)
λ 5000—λ 3900	Glass spectrograph	7 „	(mean)
	Grating—		
	1st order	5·6 „	
	3rd order	1·9 „	
λ 3900—λ 2250	Quartz Littrow spectrograph (Hilger's E. 1)	12 to 2 „	
λ 2300—λ 2100	Quartz spectrograph (Hilger's E. 2)	4 „	(mean)
λ 2100—λ 1850	Quartz spectrograph (Bellingham and Stanley)	6 „	(mean)
λ 2000—λ 850	Vacuum spectrograph—		
	1st order	17 „	
	2nd order	8·5 „	

The general catalogue which follows includes only those lines whose assignment to N III seems fairly certain. The intensities given are on an arbitrary scale of 0 to 10. The intensity scale is not uniform over long ranges but may be taken to be constant throughout each group of lines. Wave-lengths enclosed in brackets have been calculated from multiplet structures. Above λ 2000 wave-lengths are in air, below λ 2000 *in vacuo*.

Table I.—Catalogue and Classification of N III Lines.

λ.	ν.	Classification.	λ.	ν.	Classification.
6487·55 (0)	15409 9	$3p'\,{}^4P_3 - 3d'\,{}^4D_2$	4896·71 (0)	20416 2	$3p'\,{}^4D_4 - 3d'\,{}^4F_3$
6478·69 (2)	15431 0	$P_3 \quad D_3$	4884·14 (1)	20468·6	$D_4 \quad F_4$
6468·77 (00)	15454 6	$P_2 \quad D_1$	4881·81 (0)	20478·5	$D_3 \quad F_2$
6466 86 (4)	15459 2	$P_3 \quad D_4$	4873·58 (2)	20513·0	$D_3 \quad F_3$
6463·03 (2)	15468·3	$P_2 \quad D_2$	4867 18 (5)	20540·0	$D_4 \quad F_5$ $D_2 \quad F_2$
6453·95 (3)	15490·1	$P_2 \quad D_3$	4861·33 (4)	20564·8	$D_3 \quad F_4$
6450·78 (2)	15497 7	$P_1 \quad D_1$	4858·88 (3)	20575·1	$D_2 \quad F_3$
6445·05 (2)	15511 5	$P_1 \quad D_2$	4858 74 (2)	20575 7	$D_1 \quad F_2$
5314·45 (2)	18811·4	$3p'\,{}^4P_3 - 3d'\,{}^4P_3$	*4641·90 (3)	21536·8	$3p\,{}^2P_2 - 3d\,{}^2D_2$
5298·93 (1)	18866·5	$P_2 \quad P_3$	*4640·64 (10)	21542·7	$P_2 \quad D_3$
5297·86 (1)	18870 3	$P_3 \quad P_2$	*4634·16 (8)	21572·8	$P_1 \quad D_2$
5282·52 (00)	18925·1	$P_2 \quad P_2$			
5272·60 (1)	18960·7	$P_2 \quad P_1$			
5270·59 (1)	18967·9	$P_1 \quad P_2$			
5260·91 (1n)	19002·8	$P_1 \quad P_1$			

* Classified by Bowen.

Table I—(continued).

λ.	ν.	Classification.
4547·34 (0)	21984·7	$3s'\,^4P_3-3p'\,^4D_3$
4546·36 (3)	21989·4	$3p'\,^4S_2-3d'\,^4P_3$
4535·11 (2)	22044·0	S_2 P_2
4534·57 (3)	22046·6	P_3 D_3
4530·84 (1)	22064·8	P_2 D_1
4527·86 (0)	22079·3	S_2 P_1
4523·60 (4)	22100·1	P_2 D_2
4518·18 (3)	22126·6	P_1 D_1
4514·89 (7)	22142·7	P_3 D_4
4510·92 (6)	22162·2	P_2 D_3 P_1 D_2
4379·09 (10*n*)	22829·4	$4f\,^2F_{3,4}-5g\,^2G_{4,5}$
4353·66 (2)	22975·6	$3p'\,^4D_4-3d'\,^4D_3$
4348·36 (5)	23003·6	D_4 D_4
4339·52 (3)	23050·5	D_3 D_3
4335·53 (4)	23071·7	D_3 D_2
4330·44 (2)	23098·8	D_2 D_1
4330·14 (2)	23100·4	D_3 D_4
4328·15 (3)	23111·0	D_2 D_2
4323·93 (2)	23133·6	D_2 D_3 D_1 D_1
4321·37 (1)	23147·3	D_1 D_2
4294·76 (0*n*)	23278·0	
4290·80 (3*n*)	23299·2	
4290·55 (1*n*)	23300·5	
4288·72 (1*n*)	23310·5	
4288·21 (0*n*)	23313·2	
4284·51 (1*n*)	23333·4	
4200·02 (6)	23802·7	$4p\,^2P_2-5s\,^2S_1$?
4195·70 (5)	23827·2	P_1 S_1 ?
*4103·37 (9)	24363·3	$3s\,^2S_1-3p\,^2P_1$
*4097·31 (10)	24399·3	S_1 P_2
4003·64 (4*n*)	24970·2	
3998·69 (3*n*)	25001·0	
3938·52 (4)	25383·1	
3934·41 (3)	25409·6	
3792·87 (1)	26357·8	$3p'\,^4D_4-3d'\,^4P_3$
[3779·28]	[26452·9]	D_3 P_3
[3771·45]	[26507·5]	D_2 P_2
3771·08 (7)	26510·1	$3s'\,^4P_3-3p'\,^4S_2$
[3770·36]	[26515·1]	D_2 P_3
[3762·62]	[26569·7]	D_3 P_2
[3757·65]	[26604·8]	D_2 P_1
[3757·59]	[26605·2]	D_1 P_2
3754·62 (6)	26626·3	P_2 S_2
[3752·63]	[26640·3]	D_1 P_1
3745·83 (4)	26688·8	P_1 S_2

λ.	ν.	Classification.
3374·06 (6)	29629·5	$3s'\,^4P_3-3p'\,^4P_3$
3367·36 (7)	29688·4	P_2 P_2
3365·79 (3)	29702·3	P_2 P_1
3361·90 (2)	29745·0	P_3 P_2
3358·72 (1)	29764·8	P_1 P_1
3354·29 (4)	29804·4	P_2 P_3
3353·78 (4)	29808·5	P_1 P_2
3172·97 (2)	31507·1	$4p\,^2P_2-5d\,^2D_3$?
3171·14 (1)	31525·3	P_1 D_2 ?
2983·58 (6)	33507·2	
2982·07 (1)	33524·2	
2978·87 (3)	33560·2	
2977·32 (3)	33577·7	
2972·60 (4)	33631·0	
2862·26 (6*n*)	34927·3	$4f\,^2F_{3,4}-6g\,^2G_{4,5}$
2714·35 (1)	36830·3	$2p^3\,^2P_2-3d\,^2D_3$
2714·08 (3)	36834·0	P_1 D_2
2713·95 (5)	36835·8	P_2 D_2
2696·64 (1*n*)	37072·2	
2696·19 (2*n*)	37079·7	
2689·26 (4*n*)	37173·9	
2687·01 (3*n*)	37205·0	
2622·85 (2)	38115·1	
2621·19 (1)	38139·3	
2486·49 (3)	40205·2	
2484·56 (4)	40236·4	
2482·85 (1)	40264·1	
2453·85 (4)	40739·9	$3d'\,^4P_3-4p'\,^4D_4$
2459·26 (0)	40650·3	P_3 D_3
2462·56 (1)	40595·8	P_2 D_3
2463·04 (00)	40589·7	P_3 D_2
2466·24 (1)	40535·3	P_2 D_2
2468·36 (0)	40500·5	P_1 D_2
[2469·10]	[40488·5]	P_2 D_1
2471·24 (00)	40453·3	P_1 D_1
2372·46 (2)	42137·5	$3d'\,^4P_1-4p'\,^4S_2$
2370·49 (3)	42172·5	P_2 S_2
2367·43 (4)	42227·1	P_3 S_2
[2324·76]	·[43002]	$3d'\,^4P_1-4p'\,^4P_1$
2322·81 (1)	43038·1	P_2 P_1
2322·23 (0)	43048·9	P_1 P_2
2320·33 (00)	43084·2	P_2 P_2
2317·35 (0)	43139·4	P_3 P_2 P_2 P_3
2314·56 (1)	43191·6	P_3 P_3

* Classified by Bowen

Table I—(continued).

λ.	ν.	Classification.
[2274·81]	[43946·0]	$3d'\,^4D_2-4p'\,^4D_1$
2274·12 (0)	43950·3	D_1 D_1
2273·51 (1)	43971·2	D_3 D_2
2272·42 (0)	43992·3	D_2 D_2
2271·79 (0)	44004·5	D_4 D_3 D_1 D_2
2270·43 (2)	44030·8	D_3 D_3
2269·30 (0)	44052·8	D_2 D_3
2267·28 (3)	44092·0	D_4 D_4
2265·87 (0)	44119·4	D_3 D_4
2248·88 (5)	44452·8	$3d\,^2D_2-4p\,^2P_1$
2247·92 (6)	44471·7	D_3 P_2
2247·65 (2)	44477·6	D_2 P_2
2192·52 (1)	45586·2	
2191·39 (3)	45609·8	
2188·52 (3)	45669·4	
2188·62 (5)	45674·6	
2185·13 (1)	45740·2	
2151·61 (0)	46462·1	
2149·96 (0)	46497·8	$3d'\,^4D_2-4p'\,^4P_1$
[2149·4]	[46509]	D_1 P_1
2148·99 (1)	46518·8	$3d'\,^4F_2-4p'\,^4D_1$ D_2 P_2
2148·47 (3)	46529·9	F_3 D_2
2148·09 (3)	46538·2	F_4 D_3
[2147·91]	[46542]	D_2 P_2
2147·79 (2)	46544·7	D_4 P_3
[2147·32]	[46555]	D_1 P_2
2147·27 (4)	46556·0	F_5 D_4
[2146·86]	[46564·9]	F_3 D_2
2146·59 (00)	46570·8	D_3 P_1
2145·74 (1*n*)	46580·3	F_3 D_3
[2145·52]	[46594]	D_2 P_1
[2144·07]	[46625·5]	F_2 D_3
2143·96 (0)	46627·7	F_4 D_4
2142·67 (0)	46655·6	
[2141·60]	[46679·2]	F_3 D_4
2079·86 (6)	48064·9	
2074·74 (2)	48183·5	
2072·86 (1)	48227·2	
2071·79 (2)	48252·0	
2070·63 (5)	48279·1	
2068·25 (6)	48334·6	
2063·99 (10)	48434·3	
2063·50 (10)	48445·9	
2035·62 (2)	49109·1	
2035·02 (3)	49123·7	
λ vac.		
1953·80 (3)	51182·2	$3p'\,^4P_2-4s'\,^4P_1$
1953·66 (3)	51186·0	P_3 P_2
1952·20 (1)	51224·3	P_1 P_1
1951·43 (2)	51244·5	P_2 P_2
1949·81 (4)	51287·0	P_1 P_2
1949·22 (6)	51302·5	P_3 P_2
1946·99 (5)	51361·3	P_3 P_3

λ vac.	ν.	Classification.
1923·86 (2)	51978·8	
1923·11 (2)	51999·1	
1921·49 (4)	52042·9	
1920·86 (8)	52060·0	
1919·99 (2)	52083·5	
1919·71 (2)	52091·2	
1919·44 (1)	52098·5	
1919·06 (0)	52108·9	
1918·69 (0)	52118·9	
1909·12 (1)	52363·9	
1908·27 (7)	52386·2	
1907·44 (4)	52410·6	
1907·05 (1)	52418·6	
1906·38 (1)	52437·0	
1885·25 (10)	53043·3	
1845·80 (4)	54177·3	$3d\,^2D_2-4f\,^2F_{3,4}$
1845·64 (5)	54182·0	D_2 F_3
1841·68 (1)	54298	$3p'\,^4S_2-4s'\,^4P_1$
1839·59 (2)	54360	S_2 P_2
1835·58 (3)	54479	S_2 P_3
1805·5 (7)	55386	$3p\,^2P_2-4s\,^2S_1$
1804·3 (6)	55423	P_1 S_1
1751·75 (10)	57085·7	$2p'\,^2P_2-2p^3\,^2D_3$
1751·24 (6)	57102·4	P_2 D_2
1747·86 (9)	57213·0	P_1 D_2
1730·04 (8)	57802	
1699·95 (4)	58825	$3p'\,^4D_3-4s'\,^4P_2$ D_2 P_1
1699·32 (5)	58847	D_4 P_3
1699·00 (2)	58858	D_1 P_1
1698·16 (2)	58887	D_2 P_2
1697·19 (0)	58921	D_1 P_2
1696·54 (3)	58943	D_3 P_3
1694·79 (0)	59004	D_2 P_3
[1472·88]	67894]	$3p'\,^4P_3-4d'\,^4D_2$
[1472·30]	[67921]	P_3 D_3
1471·69 (2)	[67949	P_3 D_4
[1471·60]	67953]	P_2 D_2
1471·02 (1)	[67980	P_2 D_3
1470·68 (0)	67996	P_1 D_2
1387·31 (4)	72082	$2p'\,^2S_1-2p^3\,^2D_2$?
[1348·45]	[74159]	$3p'\,^4D_4-4d'\,^4F_5$
1347·56 (0)	74208	D_4 F_4
[1346·70]	[74256]	D_3 F_3
1346·27 (4)	74282	D_4 F_3
1345·69 (4)	74311	D_3 F_4 D_2 F_3

Table I—(continued).

λ vac.	ν.	Classification.
[1325·10]	[75466]	$3p'\,{}^4D_4-4d'\,{}^4D_3$
1324·40 (3)	75506	$D_4 \quad D_4$
*[1323·89]	[75535]	$D_3 \quad D_2$
[1323·42]	[75562]	$D_3 \quad D_3$
[1322·90]	[75591]	$D_3 \quad D_4$
[1322·80]	[75597]	$D_2 \quad D_2$
[1322·33]	[75624]	$D_2 \quad D_3$
[1322·17]	[75633]	$D_1 \quad D_2$
Observations and classification by Bowen		
1184·55 (3)	84420·2	$2p'\,{}^2P_2-2p^3\,{}^2P_{1,2}$
1183·04 (3)	84528·2	$P_1 \quad P_{1,2}$
1006·03 (1)	99400·6	$2p'\,{}^2S_1-2p^3\,{}^2P_{1,2}$
991·571 (7)	100850·1	$2p\,{}^2P_2-2p'\,{}^2D_{2,3}$
989·803 (7)	101030·2	$P_1 \quad D_2$
†979·9 (2)	102051	$2p'\,{}^2D_{2,3}-2p^3\,{}^2D_{2,3}$
772·980 (2)	129369·5	$2p'\,{}^2D_2-2p^3\,{}^2P_{1,2}$
772·903 (2)	129382·3	$D_3 \quad P_2$
772·384 (4)	129469·3	$2p'\,{}^4P_3-2p^3\,{}^4S_2$
771·904 (4)	129549·8	$P_2 \quad S_2$
771·545 (3)	129610·1	$P_1 \quad S_2$
764·358 (5)	130828·7	$2p\,{}^2P_2-2p'\,{}^2S_1$
763·348 (5)	131001·8	$P_1 \quad S_1$
686·340 (5)	145700·4	$2p\,{}^2P_2 \quad 2p'\,{}^2P_1$
685·820 (6)	145810·9	$P_2 \quad P_2$
685·519 (6)	145874·7	$P_1 \quad P_1$
684·997 (5)	145986·0	$P_1 \quad P_2$
452·24 (1)	221122	$2p\,{}^2P_2 \quad 3s\,{}^2S_1$
451·91 (1)	221283	$P_1 \quad S_1$
374·31 (1)	267158	$2p\,{}^2P-3d\,{}^2D$

* There is a carbon line at 1323·93

† Observed and classified by present author.

Notation.

The notation adopted for distinguishing the spectroscopic terms is similar to that recently used by Fowler ‡ In this notation the orbital designation of the "series electron" is used as a prefix to each term, a small italic letter indicating the azimuthal number of the electron orbit ($s = 1, p = 2$, etc). To distinguish between the different "families" of terms arising from different configurations of the core, undashed prefixes are employed for the principal family and dashed prefixes for the other families in accordance with the rules for electron transitions ($\Delta l = \pm 1$ for one electron and $\Delta k = 0$ or 2 for a second). Thus the ${}^2P$ terms given by the configurations :—$2(2_1)$, $1(3_2)$ and $1(2_1)$, $1(2_2)$, $1(3_1)$ will be called $3p\,{}^2P$ and $3s'\,{}^2P$ respectively. Possible combinations are indicated by applying the familiar rules for dashed and undashed terms to the prefixes. Thus, $3p\,{}^2P$ may combine with $3p'\,{}^2P$, but not with $4s'{}^2P$ or $3d'{}^2F$. On this system the terms arising from the configuration $0(2_1)$, $3(2_2)$ would have the prefix $2p''$, which would indicate that they could combine with terms having undashed prefixes such as $3s\,{}^2S$. As will appear later, the existence of such combinations is strongly suggested, but the evidence is perhaps not yet sufficient to warrant the extension of the "dashing" method to these terms. They have therefore provisionally been given the prefix $2p^3$.

---

‡ A. Fowler, 'Roy. Soc. Proc.,' A, vol. 117, p 317 (1928), and Fowler and Selwyn, 'Roy. Soc. Proc.,' A, vol. 118, p. 34 (1928).

Predicted Terms of N III

Three families of predicted terms are set out in Table II corresponding to the addition of an electron to three different states of the N IV core. The first family consists of the regular doublet terms, only one term being given by each electron configuration. In the second family, both doublet and quartet terms occur, several terms arising from a single electron configuration. The third family also contains doublet and quartet terms. Bowen* has identified the deepest row of terms in each family with the exception of $2p^3\,{}^2\mathrm{D}$. In the table, terms which have been identified are printed in heavy type.

The Doublet System.

Besides the two pairs of lines at λ 4103 and λ 4640 which, according to Fowler's suggestion, involve a $^2\mathrm{P}$ term having a wave-number separation of 36·0, another pair having a similar separation has been observed at λ 1805. Regarding λ 4103 and λ 1805 as consecutive sharp pairs—$3s\,{}^2\mathrm{S}-3p\,{}^2\mathrm{P}$ and $3p\,{}^2\mathrm{P}-4s\,{}^2\mathrm{S}$—a Rydberg formula gives $3s\,{}^2\mathrm{S} = 161730$, $4s\,{}^2\mathrm{S} = 81945$ and $3p\,{}^2\mathrm{P} = 137331$ Taking the lines at λ 4640 as $3p\,{}^2\mathrm{P}-3d\,{}^2\mathrm{D}$ thus gives $3d\,{}^2\mathrm{D} = 115789$. A comparison of these rough term values with those of C II† showed that they were of the expected order of magnitude.

	N III.	C II.	N III/C II.
$3s\,{}^2\mathrm{S}$	161730	80121	2·02
$3p\,{}^2\mathrm{P}$	137331	64923	2·12
$3d\,{}^2\mathrm{D}$	115789	51107	2·27
$4s\,{}^2\mathrm{S}$	81945	39425	2·08

By a simple Rydberg extrapolation, the values of $4p\,{}^2\mathrm{P}$ and $4d\,{}^2\mathrm{D}$ were next estimated to be about 72846 and 64242 respectively. In combination with $3d\,{}^2\mathrm{D}$, the term $4p\,{}^2\mathrm{P}$ should give an inverted diffuse doublet involving the $3d\,{}^2\mathrm{D}$ separation (5·9) at about 42943 (λ 2330). This was found at λ 2248 giving $4p\,{}^2\mathrm{P} = 71318$ ($\Delta\nu = 24{\cdot}8$). This would be expected to combine with $3s\,{}^2\mathrm{S}$ to give a doublet at λ 1106 but this has not been observed. Its combination with $4s\,{}^2\mathrm{S}$ would occur in the far infra-red, out of the range of

---

* *Loc. cit.*

† A. Fowler, 'Roy. Soc. Proc.,' A, vol. 105, p. 299 (1924).

Table II —Predicted Terms of N III

K	L		M			N					Doublet terms					Quartet terms					Term values.
1_1	2_1	2_2	3_1	3_2	3_3	4_1	4_2	4_3	4_4												
2	2	1								$2p$		$^2P^*$									383914
2	2		1							$3s$	$^2S^*$										162798
2	2			1						$3p$		$^2P^*$									138399
2	2				1					$3d$			$^2D^*$								116856
2	2					1				$4s$	^2S										83013
2	2						1			$4p$		^2P									72384
2	2							1		$4d$			^2D								
2	2								1	$4f$				^2F							62679
										$5s$	^2S										48581 ?
										$5p$		^2P									
										$5d$			^2D								40877 ?
										$5f$				^2F							
										$5g$					^2G						39849
										$6g$					^2G						27751
2	1	2								$2p'$	$^2S^*$	$^2P^*$	$^2D^*$				^4P				283058—238103
2	1	1	1							$3s'$		†^2P					^4P				‡162900
2	1	1		1						$3p'$	^2S	^2P	^2D			^4S	^4P	^4D			‡140757—133211
2	1	1			1					$3d'$		^2P	^2D	^2F			^4P	^4D	^4F		‡120217—114310
2	1	1				1				$4s'$		^2P					^4P				‡81909
2	1	1					1			$4p'$	^2S	^2P	^2D			^4S	^4P	^4D			‡73660—71208
2	1	1						1		$4d'$		^2P	^2D	^2F			^4P	^4D	^4F		‡66475—65262
2	1	1							1	$4f'$			^2D	^2F	^2G			^4D	^4F	^4G	
2		3								$2p^3$		$^{*2}P$	^2D			^4S					181017—153692
2		2	1							$3s^2$	^2S	^2P	^2D				^4P				

* Identified by Bowen.

† Except in the $2p'$ row, there are two similar sets of doublet terms in this family, only one of which is given here. One set may be considered as arising from the $2p\,^1P$ term of N IV and the other from $2p\,^3P$.

‡ Quartet terms are not correctly related to doublet terms. The former are referred to the ^3P limit of N IV, while the latter are referred to ^1S.

observations. A pair of lines at λ 4200 having a separation $\Delta\nu = 24\cdot5$ may possibly be $4p\,{}^2\mathrm{P}-5s\,{}^2\mathrm{S}$, but the considerable strength of the lines, together with the non-appearance of $3s\,{}^2\mathrm{S}-4p\,{}^2\mathrm{P}$ make it seem unlikely. The lines in question would give $5s\,{}^2\mathrm{S} = 47513\cdot9$.

Neither $4d\,{}^2\mathrm{D}$ nor $5d\,{}^2\mathrm{D}$ has been found with any degree of certainty. A pair of rather diffuse lines at λ 3172, if considered to be the combination $4p\,{}^2\mathrm{P}-5d\,{}^2\mathrm{D}$, would give 39809 as the value of $5d\,{}^2\mathrm{D}$.

An estimate of the probable value of $4f\,{}^2\mathrm{F}$ was made by multiplying the corresponding term in C II by $2\cdot25$, giving $4f\,{}^2\mathrm{F} = 62250$ This would indicate the position of $3d\,{}^2\mathrm{D}-4f\,{}^2\mathrm{F}$ as roughly 53538 (λ 1870) and a close pair of lines at λ 1846 having a separation $\Delta\nu = 4\cdot7$ probably represents this combination, giving $4f\,{}^2\mathrm{F} = 61611\cdot0$. The only other line in the vicinity, at λ 1885, is not resolvable as a doublet and so is unlikely to be a ${}^2\mathrm{D}\,{}^2\mathrm{F}$ combination

The term $5g\,{}^2\mathrm{G}$ would have a value slightly less than $5f\,{}^2\mathrm{F}$. By extrapolation this comes to about 39420, giving the approximate position of $4f\,{}^2\mathrm{F}-5g\,{}^2\mathrm{G}$ as 22191. The strong line at $22829\cdot4$ (λ 4379) is almost certainly this combination, giving $5g\,{}^2\mathrm{G} = 38781\cdot6$. As this line is so strong, it might be expected that the next member of the series, $4f\,{}^2\mathrm{F}-6g\,{}^2\mathrm{G}$, would be prominent, its estimated position being at 34890. Assuming it to be represented by a strong diffuse line at $34927\cdot3$ (λ 2862), $6g\,{}^2\mathrm{G} = 26683\cdot7$.

In order to obtain more accurate values of the terms, a formula has been calculated for the sequence of three $p\,{}^2\mathrm{P}$ terms, namely,

$$p\,{}^2\mathrm{P}_2 = 9\mathrm{R}/[m + 0\cdot73879 - 0\cdot13488/m]^2,\ \mathrm{R} = 109737,\quad m = 1, 2, 3,$$

which gives

$$2p\,{}^2\mathrm{P}_2 = 383915,\quad 3p\,{}^2\mathrm{P}_2 = 138399,\quad 4p\,{}^2\mathrm{P}_2 = 72384$$

For purposes of extrapolation, formulæ are also given below for the $s\,{}^2\mathrm{S}$ and $g\,{}^2\mathrm{G}$ terms, for each of which two successive terms have been identified.

$$s\,{}^2\mathrm{S} = 9\mathrm{R}/[m + 0\cdot42163 + 0\cdot08286/m]^2,\ m = 2, 3$$
$$g\,{}^2\mathrm{G} = 9\mathrm{R}/[m + 0\cdot91443 + 0\cdot25566/m]^2,\ m = 4, 5$$

The doublet terms of the first family are set out in Table V together with the corresponding terms of C II for comparison.

In the third family of doublet terms (see Table II) $2p^3\,{}^2\mathrm{D}$ has been identified and evaluated from its combinations with $2p'\,{}^2\mathrm{P}$, $2p'\,{}^2\mathrm{S}$ and $2p'\,{}^2\mathrm{D}$—three terms given by Bowen * The new term combines very strongly with $2p'\,{}^2\mathrm{P}$,

---

* *Loc. cit.*

giving a diffuse doublet with inverted satellite at λ 1751. Confirmation of this identification is found in the fact that this and similar combinations in C II† and O IV‡ follow the irregular doublet law very closely as is shown below :—

C II ……..	39796·4 (10)		
		− *5·1*	
	39801·5 (4)		
		41·3	17289·3
	39842·8 (8)		
N III …	57085·7 (10)		
		− *16·7*	
	57102·4 (6)		
		110·6	
	57213·0 (9)		17344·0
O IV ……	74429·7 (6)		
		− *27·2*	
	74456·9 (3)		
		244·3	
	74701·2 (5)		

Since the term $2p^3\,{}^2D$ is the only ²D term whose combinations with $2p'\,{}^2P$ should obey the irregular doublet law, this term may be regarded as firmly established. Details of its combinations with the three $2p'$ terms are given below.

Term values.	$2p^3\,{}^2D_3$ 181017·5	*−16·6*	$2p^3\,{}^2D_2$ 181000·9
$2p'\,{}^2P_2$ = 238103·2	57085·7 (10)	*−16·7*	57102·4 (6)
110·8			*110·6*
$2p'\,{}^2P_1$ = 238214·0			57213·0 (9)
$2p'\,{}^2S_1$ = 253086·4			*[72085·5] 72082 (4)
$2p'\,{}^2D_3$ = 283064·3	[102046·8]		[102063·4]
−5·7		102051 (2)	
$2p'\,{}^2D_2$ = 283058·6	[102041·1]		[102057·7]

* Calculated wave-numbers are enclosed in brackets.

---

† A. Fowler and E. W. H. Selwyn, 'Roy. Soc. Proc.,' A, vol. 120, p. 312 (1928).

‡ Author's identification.

The other $2p^3$ doublet term, namely, $2p^3\,{}^2P$, has been identified by Bowen from its combinations with the three $2p'$ terms. His value, 151822·1, becomes 153685·1 when reduced to the scale adopted in this paper. This term also appears to combine fairly strongly with $3d\,{}^2D$, giving three lines at λ 2714. As such a combination would involve a triple electron jump, it was at first thought that this indicated that the term was not $2p^3\,{}^2P$ but $3s'\,{}^2P$ which would require only a double transition. But confirmation of Bowen's identification of the term as $2p^3\,{}^2P$ is given by consideration of the irregular doublet law as applied to the differences in the values of $\sqrt{T/R}$ between C II and N III and between N III and O IV, as set out in Table III. In the case of the terms under consideration, this difference is 0·562 which shows clearly that the terms belong to the group having principal quantum number 2. If, then, the term is accepted as $2p^3\,{}^2P$, the combination $2p^3\,{}^2P-3d\,{}^2D$ requires that three electrons must make simultaneous jumps, two from 2_2 to 2_1 orbits and one from a 2_2 to a 3_3 orbit. The possibility of such triple transitions has been admitted by Epstein,[†] who suggests, however, that the resulting lines would be weak. The only other similar combination observed is a faint line at λ 1559 which is in the mean position of the lines making the combination $2p^3\,{}^2D-3d\,{}^2D$.

	$2p^3\,{}^2P_2$ *153692·0	*4·2*	$2p^3\,{}^2P_1$ 153696·2
$3d\,{}^2D_3 = 116856·3$	36835·7 (5)		
5·9	*5·4*		
$3d\,{}^2D_2 = 116862·2$	36830·3 (1)	3·7	36834·0 (3)

Some support for the suggested occurrence of triple electron jumps is given by consideration of the spectrum of C II. The term $2p^3\,{}^2P$ in C II has not been found, but Prof. Fowler has estimated its value as about 25600. Two prominent unclassified lines in C II may possibly represent the triple transition combination $3d\,{}^2D-2p^3\,{}^2P$, namely --

	$2p^3\,{}^2P_2$ 25316·3	*2·4*	$2p^3\,{}^2P_1$ 25319·2
$3d\,{}^2D_3 = 51107·6$	25791·3 (4)		
1·4			
$3d\,{}^2D_2 = 51109·0$	[25792·7]		25789·8 (2)

---

* The difference of 8 units between the value of the term given by this combination and Bowen's value may be accounted for by the difficulty of obtaining accurate wave-numbers in the far ultra-violet. At λ 750, $\Delta\nu = 8$ corresponds to $\Delta\lambda$ 0·05 A.

[†] 'Proc. Nat. Acad. of Sciences,' vol. 10, p. 341 (1924).

In Table III are given values of $\sqrt{T/R}$ and the screening constant, s, for the doublet terms of C II, N III and O IV. The screening constant values have been calculated from the formula

$$\sqrt{T/R} = \frac{1}{n}(Z - s),$$

n being the principal quantum number of the added electron, and Z the atomic number.

Table III. –Values of $\sqrt{T/R}$ and s for Doublet Terms.

	$\sqrt{T/R}$.					s		
	C II.	Diff.	N III.	Diff	O IV.	C II.	N III	O IV.
$2p\,^2P$	1·339	*0·532*	1·871	*0·512*	2·383	3·322	3·258	3·234
$2p'\,^2D$	1·053	*0·553*	1·606	*0·522*	2·128	3·894	3·788	3·744
$2p'\,^2S$	0·955	*0·564*	1·519	*0·527*	2·046	4·190	3·962	3·908
$2p'\,^2P$	0·885	*0·588*	1·473	*0·536*	2·009	4·230	4·054	3·982
$2p^3\,^2D$	0·649	*0·636*	1·285*	*0·547*	1·832*	4·602	4·430*	4·330*
$2p^3\,^2P$			1·184	*0·562*	1·746		4·632	4·508
$3s\,^2S$	0·855	*0·363*	1·218			3·435	3·346	
$3p\,^2P$	0·769	*0·354*	1·123			3·693	3·631	
$3d\,^2D$	0·683	*0·349*	1·032			3·951	3·904	

* From author's identification of term.

As would be expected, the differences in the values of $\sqrt{T/R}$ between C II and N III and between N III and O IV are roughly constant for terms having the same principal quantum number. Among the first group of terms, there appears to be some relation between the values of $\sqrt{T/R}$ and the differences, the latter increasing as the former decrease. This relation is shown in fig. 1 where $\Delta\sqrt{T/R}$ is plotted against $\sqrt{T/R}$ for N III The highest point on the upper curve corresponds to the estimated value of $2p^3\,^2P$ for C II already mentioned, and is seen to be in the position to be expected from the similar character of the two curves

Table IV is a complete list of the doublet combinations which have been identified, including Bowen's observations.

Table V contains a list of the doublet terms of N III compared with those of C II. In the last column are given the ratios of the N III terms to the corresponding C II terms. For terms having the same principal quantum number,

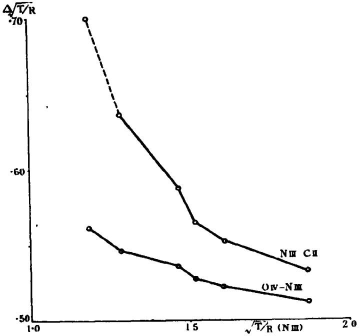


FIG. 1.

such as $4s\,^2S$, $4p\,^2P$, $4d\,^2D$ and $4f\,^2F$, this ratio apparently increases in the order given, reaching a value of about 2·25 ($3^2/2^2$) for the F terms—which would be expected to be most hydrogen-like. For the doublets of the second and third families, the ratio of the terms is considerably higher—as much as 3·92 for $2p^3\,^2D$

Table IV.—Doublet Combinations.

	$2p\,^2P_1$		$2p\,^2P_2$	$3p\,^2P_1$		$3p\,^2P_2$	$4p\,^2P_1$		$4p\,^2P_2$	$4f\,^2F_{3\,4}$
	384088·4	*174 4*	383914 0	138435 0	*36 0*	138399 0	72409 4	*24 8*	72384 6	62879 0
$3s\,^2S_1$ = 162798 3	221283 (1)	*161*	221122 (1)	24363·3 (9)	*36 0*	24399·3 (10)	[90888·9]		[90413·7]	
$3d\,^2D_2$ = 116862 2				21572 8 (8)	*36 0*	21536 8 (3)	44452·8 (5)	*24 8*	44477 6 (2)	54182·0 (5)
5 9		267158 (1)				*5 9*			*5 9*	*4 7*
$3d\,^2D_3$ = 116856·3						21542 7 (10)			44471 7 (6)	54177·3 (4)
$4s\,^2S_1$ = 83013				35423 (6)	*37*	35386 (7)				
$4d\,^2D$ = [64566]				[73869]		[73833]				
$5s\,^2S_1$ = 48581·9				[89853·1]		[89817·1]	23827 2 (5)	*24·5*	23802·7 (6)	
$5d\,^2D_2$ = 40884·1							31525·3 (1)		[31500 5]	
6·6										
$5d\,^2D_3$ = 40877·5									31507·1 (2)	
$5g\,^2G$ = 39849 6										22829 4 (10)
$6g\,^2G$ = 27751·7										34927·3 (6)

Table IV —Doublet Combinations —(continued)

	$2p\,^2P_1$		$2p\,^2P_2$	$2p^3\,^2D_3$		$2p^3\,^2D_2$	$2p^3\,^4P_3$		$2p^3\,^2P_2$
	384088 4	*1 4 4*	383914 0	181017 5	*−16 6*	181000 9	153696 2	*4 2*	153692 0
$2p$ D_3 = 283064 3								129382 3 (2)	
—5	101030 2 (7)	*180 1*	100850 1 (7)		102051 (2)			*12 8*	
$2p$ D = 283058 6								129369 5 (2)	
$2p$ $^2S_1$ — 253086 4	131001 8 (5)	*173 1*	130828 7 (5)			72082 (4)		99400 6 (1)	
$2p$ P_1 — 238214 0	145874 9 (6)	*174 5*	145700 4 (5)			57213 0 (9)		84528 2 (3)	
110 8	*111 1*		*110 5*			*110 6*		*108 0*	
$2p$ $^2P_2$ = 238103 2	145986 0 (5)	*175 1*	145810 9 (6)	57085 7 (10)	*16 7*	57102 4 (6)		84420 2 (3)	
$3d$ D_3 — 116862 2							36834 0 (3)	*3 7*	36830 3 (1)
5 9		267158 (1)			64153 (1)*				*5 5*
$3d$ D_2 = 116856 [illegible]									36835 8 (5)

Table V.—Doublet Terms of N III and C II.

	C II	Δν.	N III	Δν.	n^* for N III.	N III/C II.
$2p\,^2P_1$	196670	*58*	384088·4†	*174·4*		
$2p\,^2P_2$	196612		383914·0		1·603	1·95
$3s\,^2S_1$	80121		162798·3†		2·463	2·03
$3p\,^2P_1$	64934	*11*	138435·0†	*36·0*		
$3p\,^2P_2$	64923		138399·0		2·671	2·13
$3d\,^2D_2$	51108·9	*1·3*	116862·2†	*5·9*		
$3d\,^2D_3$	51107·6		116856·3		2·906	2·29
$4s\,^2S_1$	39425		83013		3·440	2·11
$4p\,^2P_1$	34140	*6*	72409·4	*24·8*		
$4p\,^2P_2$	34134		72384·6		3·694	2·12
$4d\,^2D_2$	28535·1	*0·4*				
$4d\,^2D_3$	28534·7		[64566]		[3·910]	
$4f\,^2F$	27680		62679·0		3·969	2·26
$5s\,^2S_1$	23311		[50046]		[4·442]	
$5p\,^2P_1$						
$5p\,^2P_2$			[44595]		[4·705]	
$5d\,^2D_2$						
$5d\,^2D_3$	18164		40877·5 ?		4·914 ?	2·25 ?
$5f\,^2F$	17703		[39960]			
$5g\,^2G$			39849·6		4·978	
$6g\,^2G$			27751·7		5·965	
$2p'\,^2D_3$			283064·3†	*−5·7*		
$2p'\,^2D_2$	121727		283058·6		1·867	2·32
$2p'\,^2S_1$	100165		253098·4†		1·975	2·53
$2p'\,^2P_1$	86033·9	*41·6*	238214·0†	*110·8*		
$2p'\,^2P_2$	85992·3		238103·2		2·036	2·77

Table V- (continued).

	C II	$\Delta\nu$.	N III	$\Delta\nu$.	n^* for N III.	N III/C II
$2p^3\,{}^2P_1$			153696·2	*4·2*		
$2p^3\,{}^2P_2$			153692·0		2·534	
$2p^3\,{}^2D_3$	46196·2	*−5·1*	181017·5	*−16·6*		3·92
$2p^3\,{}^2D_2$	46191·1		181000·9		2·336	

† Bowen's values for these terms have been increased by 1863 units to bring them to the scale adopted in this paper.

The Quartet System.

A beginning was made in the identification of the quartet system by the recognition of an isolated group of seven lines at λ 1950 as a PP′ multiplet involving terms with wave-number separations of 116·5, 62·7 and 58·9, 43·5 respectively. Another group of seven lines at λ 3350 was found to have almost identical separations (116·1, 62·5 and 58·8, 43·5). It appeared that these two groups involved three ${}^4P_{321}$ terms of which two had the same separations and would probably be successive members of a sequence. From consideration of the predicted terms (Table II) it seemed likely that $3s'\,{}^4P$ and $4s'\,{}^4P$ were the terms in question, since Bowen had already found different separations of 80, 60 for $2p'\,{}^4P$ The two similar groups were therefore assumed to be $3s'\,{}^4P—3p'\,{}^4P$ and $3p'\,{}^4P—4s'\,{}^4P$. To complete the triad of $3s'—3p'$ combinations, a strong ${}^4P—{}^4S$ triplet was found at λ 3771 with separations of 116·2, 62·5 and a ${}^4P—{}^4D$ multiplet at λ 4514 which gave the same ${}^4P$ separations (116·2, 62·5) and also a ${}^4D$ term having separations of 96·2, 62·2 35·5

It was now possible to calculate the positions in which the remaining two $3p'—4s'$ combinations should occur. They were both found, $3p'\,{}^4D—4s'\,{}^4P$ at λ 1700 and $3p'\,{}^4S—4s'\,{}^4P$ at λ 1835. The six multiplets identified up to this point are set out below, the term values being those finally adopted.

Term values.		$3s'\,{}^4P_3$		$3s'\,{}^4P_2$		$3s'\,{}^4P_1$
		162900·0	115·8	163015·8	62·5	163078·3
$3p'\,{}^4P_3$ =	133211·6	29688·4 (7)	116·0	29804·4 (4)		
	58·9	58·9		59·4		
$3p'\,{}^4P_2$ =	133270·5	29629·5 (6)	115·5	29745·0 (2)	63·5	29808·5 (4)
	43·5			42·7		43·7
$3p'\,{}^4P_1$ =	133314·0			29702·3 (3)	62·5	29764·8 (1)
$3p'\,{}^4D_4$ —	140757·2	22142·8 (7)				
	96·2	96·2				
$3p'\,{}^4D_3$ =	140853·4	22046·6 (3)	115·7	*22162·3 (6)		
	62·2	61·9		62·2		
$3p'\,{}^4D_2$ =	140915·6	21984·7 (0)	115·4	22100·1 (4)	62·2	*22162·3 (6)
	35·5			35·3		35·5
$3p'\,{}^4D_1$ =	140951·1			22064·8 (1)	62·0	22126·8 (3)
$3p'\,{}^4S_2$ =	136389·9	26510·1 (7)	116·2	26626·3 (6)	62·5	26688·8 (4)

Term values.		$4s'\,{}^4P_3$		$4s'\,{}^4P_2$		$4s'\,{}^4P_1$
		81909·1	116·5	82025·6	62·7	82088·3
$3p'\,{}^4P_3$ =	133211·6	51302·5 (6)	116·5	51186·0 (3)		
	58·9	58·8		58·5		
$3p'\,{}^4P_2$ =	133270·5	51361·3 (5)	116·8	51244·5 (2)	62·3	51182·2 (3)
	43·5			42·5		42·1
$3p'\,{}^4P_1$ =	133314·0			51287·0 (4)	62·7	51224·3 (1)
$3p'\,{}^4D_4$ =	140757·2	58847 (5)				
	96·2	96				
$3p'\,{}^4D_3$ =	140853·4	58943 (3)	118	*58825 (4)		
	62·2	61		62		
$3p'\,{}^4D_2$ =	140915·6	59004 (0)	117	58887 (2)	62	*58825 (4)
	35·5			34		33
$3p'\,{}^4D_1$ =	140951·1			58921 (0)	63	58858 (2)
$3p'\,{}^4S_2$ =	136389·9	54479 (3)	119	54360 (2)	62	54298 (1)

* Used twice

In searching for combinations involving the next ($3d'$) row of terms, two triplets—one at λ 2370, the other at λ 4546—seemed to be significant. They have identical separations (54·6, 35·1) and, from their positions (one inverted with respect to the other), appeared likely to represent the combinations $3p'\,{}^4S - 3d'\,{}^4P$ and $3d'\,{}^4P - 4p'\,{}^4S$. The value of $3p'\,{}^4S$ being known, $4p'\,{}^4S$ and $3d'\,{}^4P$ were worked out on this assumption, the values given being, $4p'\,{}^4S = 72172$, $3d'\,{}^4P = 94162$. This value for $4p'\,{}^4S$ is in good Rydberg sequence with $3p'\,{}^4S$ (136389). The term $3d'\,{}^4P$ should combine with $3p'\,{}^4P$ and $3p'\,{}^4D$, but no such combinations were found in the positions indicated. However, on trying the triplets the other way round (that is, assuming an inverted

$3d'\,{}^4P$ term equal to 114400) the expected combinations were found in their correct positions, $3p'\,{}^4P-3d'\,{}^4P$ at λ 5320 and $3p'\,{}^4D-3d'\,{}^4P$ at λ 3770. The latter multiplet is very weak, only the strongest line showing; the other lines are hidden by the strong triplet $3s'\,{}^4P-3p'\,{}^4S$ and by impurity lines due to oxygen. Details of these are given later.

Of the other two $3d'$ terms, the ${}^4D$ term was found to combine with $3p'\,{}^4D$ and with $3p'\,{}^4P$ giving groups at λ 4345 and λ 6467. A fairly strong multiplet at λ 4867 proved to be a combination between the remaining ${}^4F$ term of this row and $3p'\,{}^4D$. Combinations between a new ${}^4D$ term of the $4p'$ row with $3d'\,{}^4F$, $3d'\,{}^4D$ and $3d'\,{}^4P$ accounted for groups of lines at λ 2147, λ 2270 and λ 2453 respectively. Some difficulty was experienced in identifying the ${}^4P$ term of the $4p'$ row, but its combinations with $3d'\,{}^4P$ and $3d'\,{}^4D$ were eventually found at λ 2315 and λ 2147. The former group is intermingled with some strong N II lines, and the latter with the N III group $3d'\,{}^4F-4p'\,{}^4D$.

By a Rydberg extrapolation, the $4d'$ terms were estimated to have values between 62000 and 67000. These should combine with the $3p'$ terms to give groups of lines in the Schumann region. A narrow triplet at λ 1475 is probably made up of the strong lines of the combination $3p'\,{}^4P-4d'\,{}^4D$ and a line at λ 1324 would then fall in the position of the strongest line of $3p'\,{}^4D-4d'\,{}^4D$. A somewhat nebulous line at λ 1346 may be part of $3p'\,{}^4D-4d'\,{}^4F$, but the absence of faint lines makes it impossible to confirm this identification.

The additional multiplets thus traced are given in the table below.

Term values.	$4p'\,{}^4D_4$		$4p'\,{}^4D_3$		$4p'\,{}^4D_2$		$4p'\,{}^4D_1$
	73660·6	*89·5*	73750·1	*60·5*	73810·6	*46·7*	73857·3
$3d'\,{}^4F_5$ = 120217·2	46556·0 (4)						
71·4	*71·7*						
$3d'\,{}^4F_4$ = 120288·6	46627·7 (0)	*89·5*	46538·2 (3)				
51·8			*51·1*				
$3d'\,{}^4F_3$ = 120340·4	[46679·8]		46589·3 (1*n*)	*59·4*	46529·9 (3)		
35·1							
$3d'\,{}^4F_2$ = 120375·5			[46625·4]		[46564·9]		46518·8 (1)
$3d'\,{}^4D_4$ = 117753·6	44092·0 (3)	*87·5*	44004·5 (0)*				
28·3	*27·4*		*26·3*				
$3d'\,{}^4D_3$ = 117781·9	44119·4 (0)	*88·6*	44030·8 (2)	*59·6*	43971·2 (1)		
22·0			*22·0*		*21·1*		
$3d'\,{}^4D_2$ = 117803·9			44052·8 (0)	*60·5*	43992·3 (0)		[43946·0]
13·4					*12·2*		
$3d'\,{}^4D_1$ = 117817·3					44004·5 (0)*	*45·2*	43959·3 (0)
$3d'\,{}^4P_3$ = 114400·5	40739·9 (4)	*89·6*	40650·3 (0)	*60·6*	40589·7 (00)		
−54·6			*−54·5*		*−54·4*		
$3d'\,{}^4P_2$ = 114345·9			40595·8 (1)	*60·5*	40535·3 (1*b*)		[40488·8]
−35·1					*−34·8*		
$3d'\,{}^4P_1$ = 114310·8					40500·5 (0)	*47·2*	40453·3 (00)

* Used twice.

Term values.		$3d'\,{}^4P_3$		$3d'\,{}^4P_2$		$3d'\,{}^4P_1$
		114400·5	*−54·6*	114345·9	*−35·1*	114310·8
$3p'\,{}^4S_2$	136389·9	21989·4 (3)	*−54·6*	22044·0 (2)	*−35·3*	22079·3 (0)
$4p'\,{}^4S_2$	− 72173·4	42227·1 (4)	*54·6*	42172·5 (3)	*−35·0*	42137·5 (2)
$3p'\,{}^4P_3$	− 133211·6	18811·4 (2)	*−55·1*	18866·5 (1)		
	58·9	*58·9*		*58·6*		
$3p'\,{}^4P_2$	− 133270·5	18870·3 (1)	*−54·8*	18925·1 (00)	*− 35·6*	18960·7 (1)
	43·5			*42·8*		*42·1*
$3p'\,{}^4P_1$	− 133314·0			18967·9 (1)	*−34·9*	19002·8 (1*n*)
$3p'\,{}^4D_4$	= 140757·2	26358 (1)				
	96·2					
$3p'\,{}^4D_3$	140853·4	[26452·9]		[26507·5]*		
	62·2					
$3p'\,{}^4D_2$	− 140915·6	[26515·1]*		[26569·7]†		[26604·8]‡
	35·5					
$3p'\,{}^4D_1$	− 140951·1			[26605·2]‡		[26640·3]

* Hidden by N III 26510·1 (7)
† Hidden by O II 26569·6 (5*r*)
‡ Hidden by O III 26607·9 (5).

Term values.		$3d'\,{}^4D_4$		$3d'\,{}^4D_3$		$3d'\,{}^4D_2$		$3d'\,{}^4D_1$
		117753·6	*28·3*	117781·9	*22·0*	117803·9	*13·4*	117817·3
$3p'\,{}^4D_4$	− 140757·2	23003·6 (5)	*28·0*	22975·6 (2)				
	96·2	*96·8*		*96·1*				
$3p'\,{}^4D_3$	− 140853·4	23100·4 (2)	*28·7*	23071·7 (4)	*21·2*	23050·5 (3)		
	62·2			*61·9*		*60·5*		
$3p'\,{}^4D_2$	− 140915·6			23133·6 (2)*	*22·6*	23111·0 (3)	*12·2*	23098·8 (2)
	35·5					*36·3*		*34·8*
$3p'\,{}^4D_1$	140951·1					23417·3 (1)	*13·7*	23133·6 (2)*
$3p'\,{}^4P_3$	− 133211·6	15459·2 (4)	*28·2*	15431·0 (2)	*21·1*	15409·9 (0)		
	58·9			*59·1*		*58·4*		
$3p'\,{}^4P_2$	− 133270·5			15490·1 (3)	*21·8*	15468·3 (2)	*13·7*	15454·6 (00)
	43·5					*43·2*		*43·1*
$3p'\,{}^4P_1$	−− 133314·0					15511·5 (2)	*13·8*	15497·7 (2)

* Used twice

Term values.		$3p'\,{}^4D_4$		$3p'\,{}^4D_3$		$3p'\,{}^4D_2$		$3p'\,{}^4D_1$
		140757·2	*96·2*	140853·4	*62·2*	140915·6	*35·5*	140951·1
$3d'\,{}^4F_5$	−− 120217·2	20540·0 (5)*						
	71·4	*71·4*						
$3d'\,{}^4F_4$	− 120288·6	20468·6 (1)	*96·2*	20564·8 (4)				
	51·8	*52·4*		*51·8*				
$3d'\,{}^4F_3$	= 120340·4	20416·2 (0)	*96·8*	20513·0 (2)	*62·1*	20575·1 (3)		
	35·1			*34·5*		*35·1*		
$3d'\,{}^4F_2$	= 120375·5			20478·5 (0)	*61·5*	20540·0 (5)*	*35·7*	20575·7 (2)

* Used twice.

Term values.	$4p'\,{}^4P_3$		$4p'\,{}^4P_2$		$4p'\,{}^4P_1$
	71208·9	*52·9*	71261·8	*44·8*	71306·6
$3d'\,{}^4P_3$ = 114400·5	43191·6 (1)	*52·2*	43139·4 (0)*		
−54·6	*−52·2*		*−55·2*		
$3d'\,{}^4P_2$ = 114345·9	43139·4 (0)*	*55·2*	43084·2 (00)	*46·1*	43038·1 (1)
−35·1			*−35·3*		
$3d'\,{}^4P_1$ = 114310·8			43048·9 (0)		[43002]
$3d'\,{}^4D_4$ = 117753·6	46544·7 (2)				
28 7	*26·1*				
$3d'\,{}^4D_3$ = 117781·3	46570·8 (00)	*52·0*	46518·8 (1)		
22·0					
$3d'\,{}^4D_2$ = 117803·3	[46594]		[46542]		46497·8 (0)
13·7					
$3d'\,{}^4D_1$ = 117817·0			[46555]		[46509]

Term values.	$4d'\,{}^4D_4$		$4d'\,{}^4D_3$		$4d'\,{}^4D_2$	$4d'\,{}^4D_1$
	65262	*29*	65291	*27*	65318	——
$3p'\,{}^4P_3$ = 133211·6	67949 (2)		[67921]		[67894]	
58·9						
$3p'\,{}^4P_2$ = 133270·5			67980 (1)		[67953]	——
43·5						
$3p'\,{}^4P_1$ = 133314·0					67996 (0)	——
$3p'\,{}^4D_4$ = 140757·2	75506 (3)					

Term values.	$4d'\,{}^4F_5$		$4d'\,{}^4F_4$		$4d'\,{}^4F_3$	$4d'\,{}^4F_2$
	66475	*74*	66549	*49*	66598	——
$3p'\,{}^4D_4$ = 140757·2	74282 (4)	*74*	74208 (0)		[74159]	
96 2			*103*			——
$3p'\,{}^4D_3$ = 140853·4			74311 (4)		[74256]	
62·2						——
$3p'\,{}^4D_2$ = 140915·6					74311 (4)	
35·5						——
$3p'\,{}^4D_1$ = 140951·1						

* Used twice.

The deepest of the quartet terms, $2p'\,{}^4P$ has not yet been evaluated. Bowen has given a triplet at λ 772 as being $2p'\,{}^4P-2p^3\,{}^4S$, but has not suggested any other combinations involving either of these terms. $2p'\,{}^4P$ would be expected to combine strongly with $3s'\,{}^4P$. By analogy with the doublet terms, it seems probable that the value of $2p'\,{}^4P$ is about 380000, giving the combination $2p'\,{}^4P-3s'\,{}^4P$ at about λ 465. This estimate is necessarily very rough and may be in error by a hundred or so Ångströms.

In order to exhibit clearly the quartet combinations which have been observed, they have been summarised in Table VI, where only the strongest component of each term, and the strongest line of each multiplet are tabulated.

Table VI.—Summary of Quartet Combinations Observed.

		$3p'\,{}^4\mathrm{S}$	$3p'\,{}^4\mathrm{P}$	$3p'\,{}^4\mathrm{D}$	$4p'\,{}^4\mathrm{S}$	$4p'\,{}^4\mathrm{P}$	$4p'\,{}^4\mathrm{D}$
		136389	133211	140757	72173	71208	73660
$3s'\,{}^4\mathrm{P}$	162900	26510 (7)	29688 (7)	22142 (7)	[90727]	[91692]	[89239]
$3d'\,{}^4\mathrm{P}$	114400	21989 (3)	18811 (2)	26358 (1)	42227 (4)	43191 (1)	40739 (4)
$3d'\,{}^4\mathrm{D}$	117753		15457 (5)	23003 (5)		46544 (2)	44092 (3)
$3d'\,{}^4\mathrm{F}$	120217			20540 (5)			46556 (4)
$4s'\,{}^4\mathrm{P}$	81909	54479 (3)	51302 (6)	58847 (5)	infra-red	infra-red	infra-red
$4d'\,{}^4\mathrm{D}$	65262		67949 (2)	75506 (3)		infra-red	infra-red
$4d'\,{}^4\mathrm{F}$	66475			74282 (4)			infra-red

The only possible combinations occurring in accessible regions of the spectrum which have not been observed are those between $3s'\,{}^4\mathrm{P}$ and the $4p'$ terms. They should give groups near λ 1100 but no trace of them has been found. It is interesting to note that the corresponding doublet combination, $3s\,{}^2\mathrm{S}$—$4p\,{}^2\mathrm{P}$, which should give a pair at λ 1106 is also missing or abnormally faint.

Quartet Term Values.

For determining the values of the quartet terms, there are six pairs of successive terms for each of which a Rydberg formula may be computed. The values adopted in Table VIII are such as to divide up more or less equally the slight divergences from Rydberg sequences. As may be seen from Table VII, none of the sequences depart very much from Rydberg formulæ.

Table VII.—n^* values for Successive Quartet Terms.

$s'\,{}^4\mathrm{P}$	2·462	3 468	$d'\,{}^4\mathrm{F}$	2 865	3·853
$p'\,{}^4\mathrm{D}$	2 648	3 661	$d'\,{}^4\mathrm{D}$	2 896	3 889
$p'\,{}^4\mathrm{S}$	2 690	3 697	$d'\,{}^4\mathrm{P}$	2 938	
$p'\,{}^4\mathrm{P}$	2 722	3 723			

No intercombinations between doublet and quartet terms have been found, but some idea of their correct relative values may be obtained by considering the respective states of the N IV core from which the two sets of terms arise.

All the doublet term values are based on Rydberg sequences in the first family of terms (see Table II). These are referred to the deepest term, ^1S, of N IV. The quartet term values, on the other hand, are derived from Rydberg sequences in the second family of terms, which are referred to the second deepest term, ^3P, of N IV.

The two sets of terms, independently determined, are therefore not correctly related to each other, but are referred to different limits or different zeros; the zero of the doublet terms being deeper than that of the quartet terms by an amount equal to $^1S-^3P$ for N IV. $^1S_0-^3P_1$ is the resonance line, which has been found for B II and C III by Bowen. Application of the irregular doublet law suggests that the resonance line for N IV should occur about 58100. There is a very strong, easily excited line of N IV at 58187 (λ 1718·6) which, therefore, presumably is the resonance line $^1S_0-^3P_1$. Hence, it appears probable that the quartet term values for N III should be decreased by about 58000 to bring them to the same zero as the doublet terms.

While the great intensity of the lines above suggested as the resonance lines of B II, C III and N IV throws some doubt on the identifications,* it is probable that the resonance lines cannot be very far from these positions. This follows from a comparison of the values of the doublet terms with those of the quartet terms such as has been made for C II by Fowler and Selwyn. Thus, for N III, $3p'\ ^4P$ (133314) and $4p'\ ^4P_1$ (71307), referred to ^3P of N IV, are rather smaller than $3p\ ^2P$ (138435) and $4p\ ^2P$ (72409), referred to ^1S of N IV. It may therefore be inferred that $2p'\ ^4P_1$ referred to ^3P of N IV, is not very much smaller than $2p\ ^2P_1$ (384088), referred to ^1S of N IV. So we may write :—

$$2p'\ ^4P_1 < 384088 \qquad (^3P \text{ limit}). \qquad (1)$$

Now in the $2p^3$ row of terms, the deepest, by Hund's rule would be $2p^3\ ^4S$. Therefore

$$2p^3\ ^4S > 181018\ (2p^3\ ^2D_3) \qquad (^1S \text{ limit}).$$

But according to Bowen's identification,

$$2p'\ ^4P - 2p^3\ ^4S = 129469.$$

Therefore

$$2p'\ ^4P > 129469 + 181018.$$

Therefore

$$2p'\ ^4P > 310487 \qquad (^1S \text{ limit}). \qquad (2)$$

---

* *Cf.* Fowler and Selwyn.

Combining (1) and (2),

$$(^1\text{S limit})-(^3\text{P limit}) < 384088 - > 310487 = < 73601.$$

It therefore seems probable that the resonance line, $^1S_0-^3P_1$, for N IV is on the red side of λ 1360 and not far from the strong line at λ 1718·6 mentioned above.

In Table VIII the quartet terms for N III are collected and compared with those for C II. The similarity between the two sets of terms is well shown by comparing the values of n^*. These are plotted for C II and N III side by side in fig. 2. Among the terms of principal quantum number 3, it is noticeable

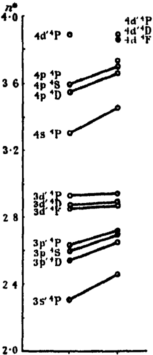


FIG. 2.

that the values of n^* for the two spectra become more nearly equal as the azimuthal quantum number increases. The same tendency is shown among the terms of principal quantum number 4, although, since these have not all been identified, the comparison is not complete. Another striking similarity is brought out by the ratios of the term separations, given in the last column of Table VIII.

Table VIII.—Quartet Terms of N III and C II.

	C II.	$\Delta\nu$.	n^*.	N III.	$\Delta\nu$.	n^*.	N III/C II.	
							$\Delta\nu$.	$\Sigma\Delta\nu$.
$3s'\ {}^4P_1$	82850·0			163078·3				
		23·8			*62·5*		2·63	
${}^4P_2$	82826·2			163015·8				2·60
		44·9			*115·8*		2·58	
${}^4P_3$	82781·3		2·302	162900·0		2·462		
$3p'\ {}^4D_1$	68120·2			140951·1				
		14·7			*35·5*		2·42	
${}^4D_2$	68105·5			140915·6				
		25·0			*62·2*		2·49	2·55
${}^4D_3$	68080·5			140853·4				
		36·3			*96·2*		2·65	
${}^4D_4$	68044·2		2·539	140757·2		2·648		
$3p'\ {}^4S_2$	65126·0		2·506	136389·9		2·690		
$3p'\ {}^4P_1$	63389·7			133314·0				
		16·3			*43·5*		2·67	
${}^4P_2$	63373·4			133270·5				2·65
		22·4			*58·9*		2·63	
${}^4P_3$	63350·9		2·632	133211·6		2·722		
$3d'\ {}^4F_2$	54063·9			120375·5				
		14·3			*35·1*		2·45	
${}^4F_3$	54049·6			120340·4				
		19·6			*51·8*		2·65	2·58
${}^4F_4$	54030·0			120288·6				
		27·6			*71·4*		2·50	
${}^4F_5$	54002·4		2·850	120217·2		2·865		
$3d'\ {}^4D_1$	53258·5			117817·3				
		5·6			*13·4*		2·40	
${}^4D_2$	53252·9			117803·9				
		8·7			*22·0*		2·53	2·59
${}^4D_3$	53244·2			117781·9				
		10·3			*28·3*		2·75	
${}^4D_4$	53233·9		2·871	117753·6		2·896		
$3d'\ {}^4P_3$	50972·7		2·934	114400·5		2·938		
		−21·5			*−54·6*		2·54	
${}^4P_2$	50951·2			114345·9				2·51
		−14·2			*−35·1*		2·47	
${}^4P_1$	50937·0			114310·8				
$4s'\ {}^4P_1$	40264·4			82088·3				
		24·0			*62·7*		2·61	
${}^4P_2$	40240·4			82025·6				2·56
		46·1			*116·5*		2·53	
${}^4P_3$	40194·3		3·304	81909·1		3·468		

Table VIII—(continued).

	C II.	$\Delta\nu$.	n^*.	N III.	$\Delta\nu$.	n^*.	N III/C II.	
							$\Delta\nu$.	$\Sigma\Delta\nu$.
$4p'\,{}^4D_1$	35056·4	*14·3*		73857·3	*46 7*		3·26	2·82
${}^4D_2$	35042·1	*22 0*		73810 6	*60 5*		2·75	
${}^4D_3$	35020 1	*33·4*		73750·1	*89 5*		2 68	
${}^4D_4$	34986·7		3·541	73660·6		3·661		
$4p'\,{}^4S_2$	34049 1		3·590	72173 4		3·697		
$4p'\,{}^4P_1$				71306 6	*44·8*			
${}^4P_2$				71261 8	*52·9*			
${}^4P_3$				71208·9		3·723		
$4d'\,{}^4F_2$				——				
${}^4F_3$				66598	*49*			
${}^4F_4$				66549	*74*			
${}^4F_5$				66475		3 853		
$4d'\,{}^4D_1$				——				
${}^4D_2$				65318	*27*			
${}^4D_3$				65291	*29*			
${}^4D_4$				65262		3·889		
$4d'\,{}^4P_3$	29006·2	*−20·5*	3·889	[63700]	[*−56*]	[3·936]		
${}^4P_2$	28985 7	*−11 9*			[*−33*]			
${}^4P_1$	28973·8							

Table IX contains values of $\sqrt{T/R}$ and the screening constant s for the quartet terms of N III and C II. The table exhibits the approximate constancy of the differences in the values of $\sqrt{T/R}$ between C II and N III for terms having the same principal quantum number, and generally confirms the identifications.

Table IX.—Values of $\sqrt{T/R}$ and s for Quartet Terms.

	$\sqrt{T/R}$.			s.	
	C II.	Difference.	N III.	C II.	N III.
$3s'\ {}^4\mathrm{P}$	0·869	*0·349*	1·218	3·393	3·346
$3p'\ {}^4\mathrm{D}$	0·788	*0·345*	1·133	3·636	3·601
$3p'\ {}^4\mathrm{S}$	0·771	*0·344*	1·115	3·687	3·655
$3p'\ {}^4\mathrm{P}$	0·760	*0·342*	1·102	3·720	3·694
$3d'\ {}^4\mathrm{F}$	0·702	*0·345*	1·047	3·894	3·859
$3d'\ {}^4\mathrm{D}$	0·697	*0·339*	1·036	3·909	3·892
$3d'\ {}^4\mathrm{P}$	0·682	*0·339*	1·021	3·954	3·937
$4s'\ {}^4\mathrm{P}$	0·605	*0·259*	0·864	3·580	3·544
$4p'\ {}^4\mathrm{D}$	0·565	*0·254*	0·819	3·740	3·724
$4p'\ {}^4\mathrm{S}$	0·557	*0·254*	0·811	3·772	3·756

Summary.

The spectrum of doubly-ionised nitrogen (N III) has been further investigated over a range extending from the far red to the extreme ultra-violet and some 170 previously unrecorded lines have been measured. In all, about 120 lines have been newly classified, 100 in the quartet system and 20 in the doublet system.

The author wishes to express his deep indebtednes to Prof. Fowler for permission to use his results for C II prior to publication, and for his continual guidance and encouragement which have been such an inspiration in preparing this paper.

The Sorption of Carbon Tetrachloride at Low Pressures by Activated Charcoals. Part I.—Apparatus and Method.

By Rufus Chaplin.

(Communicated by F. G. Donnan, F.R.S.—Received July 30, 1928.)

(1) *Introductory.*

The term "sorption," as used in this paper, has the significance given to it by McBain,[*] that is to say, it denotes simply the taking up of substances by charcoal without reference to the nature of the process.

The few data available on the subject of sorption at low pressures show that it is in this pressure region that the best promise of the establishment of a general sorption theory lies; for certain effects which occur simultaneously at higher pressures may be observed separately at low pressures, which greatly simplifies the problem. On the practical side a knowledge of the capacities of different charcoals for absorbing traces of impurities, for producing a high vacuum, etc., is most important. The practical applications of charcoal have, in fact, outstripped not only theory, but also practical methods for assessing the relevant efficiencies of different charcoals.

Some interesting data referring to the sorption of certain gases (notably oxygen, nitrogen, hydrogen, carbon dioxide and carbon monoxide) at low pressures by activated charcoal have been obtained by various workers[†] since the investigations of Dewar on the production of high vacua by the use of cooled charcoal; but, with the exception of some work by Coolidge,[‡] there are so far no data available referring to the sorption of any vapour at low pressures. This is probably due partly to the difficulty of accurately measuring low pressures of readily condensible vapours, and partly to the fact that the conditions for obtaining true equilibrium data at low pressures have not been found.

In the present paper the author describes a method of investigating in absence of foreign gases the sorption of carbon tetrachloride at low pressures, and defines certain conditions which must be fulfilled in order to obtain results representing true equilibrium.

---

[*] 'Phil. Mag.,' vol. 18, p. 916 (1909).

[†] Claude, 'C. R.,' vol. 158, p. 861 (1914); Chaplin, 'Phil. Mag.,' vol. 2, p. 1198 (1926); Magnus and Cahn, 'Z. Anorg. Chem.,' vol. 155, p. 204 (1926); Rowe, 'Phil. Mag.,' vol. 1, pp. 109, 1042 (1926).

[‡] 'J. Amer. Chem. Soc.,' vol. 46, p. 596 (1924).

The investigation comprises the determination of isotherms (quantity sorbed expressed as function of pressure at constant temperature) and of isosteres (pressure expressed as function of temperature at constant quantity) for carbon tetrachloride and charcoal. The isotherms were determined at 25° C. within the pressure limits 1×10^{-4} and $2 \cdot 3 \times 10^{-1}$ mm. of mercury. The isosteres for several definite charges of carbon tetrachloride were determined within the same pressure limits and between 0° and 70° C. Six activated charcoals of varying origin were investigated and the effect of temperature of evacuation, etc., studied. The measured magnitudes which the investigation involves are (*a*) the quantity of carbon tetrachloride taken up by the charcoal, (*b*) its pressure at equilibrium, (*c*) the temperature.

Measurement (*a*) was carried out by direct weighing, the charcoal being contained in a small counterpoised vessel. The weighings corresponded to an accuracy of $0 \cdot 1$ milligram of carbon tetrachloride per gram of charcoal.

Measurement (*b*) was effected by means of the Pirani gauge, which will shortly be described. These pressure readings constituted the greatest difficulty of the investigation owing to the persistence of traces of foreign substances and of gases displaced from the charcoal.

Temperatures were read on a standard thermometer to within $0 \cdot 1°$ C.

(2) *The Pirani Gauge.*

This is a gauge for measuring gas or vapour pressures below the order 10^{-1} mm. of mercury. The principle of the use of the gauge as developed by Campbell[*] has been fully explained by him and depends on the cooling of an electrically heated filament by the surrounding gas or vapour. The higher the gas or vapour pressure the greater is the rate of cooling. In Campbell's method of using the gauge the filament temperature is kept constant throughout and the voltage required for this is measured on a sensitive voltmeter. Under these conditions the relation between the pressure and voltage reading is

$$p = \alpha \, (v^2 - v_0^2)/v_0^2,$$

where v = voltage reading for pressure p ; v_0 = voltage reading for $p = 0$; α = a constant characteristic of the gas or vapour employed.

The voltage function $(v^2 - v_0^2)/v_0^2$ is more briefly written $f(v)$.

The linear relation between p and $f(v)$ only holds up to pressures of the order 10^{-1} mm. of mercury, the exact pressure depending on the substance under investigation In the case of carbon tetrachloride pressures up to

---

[*] 'Proc. Phys. Soc.,' vol. 32, p. 287 (1921).

0·3 mm. may be safely measured on the Pirani gauge, using the above relationship between p and $f(v)$.

Sensitivity of the Gauge.—The sensitivity of the gauge appears to be limited only by that of the voltmeter employed, this point having been specially investigated. Two voltmeters were employed in the present work; one reading from 3·500 to 5·000 volts and sensitive to 0·001 volt; and one reading from 3·5 to 15·0 volts and sensitive to 0·01 volt. These instruments were used in conjunction with a pointer galvanometer reading 0·25 microamperes per millimetre of scale. For pressures below $2 \cdot 5 \times 10^{-2}$ mm. the more sensitive instrument was used and this responded to pressures changes as small as $2 \cdot 5 \times 10^{-5}$ mm. For pressures above $2 \cdot 5 \times 10^{-2}$ mm. the other voltmeter was used, the sensitivity of which was $2 \cdot 5 \times 10^{-4}$ mm. The use of the more sensitive instrument was only made possible by taking certain special measures to suppress extraneous vapour pressures (*q.v.*).

Calibration of the Gauge.—In order to determine the value and range of constancy of α, the Pirani gauge must be calibrated against known pressures of the substance for which it is to be employed. The McLeod gauge was found to give very discordant results with low pressures of carbon tetrachloride vapour, and another method was employed. Use was made of the vapour pressure/temperature curve for carbon tetrachloride between − 22·9° C. (the melting point) and − 117° C. Between − 22·9° C. and − 50° C. the vapour pressures are high enough to be read accurately on a U mercury manometer, using a cathetometer, and between − 63° C. and − 117° C. the vapour pressures are low enough for the corresponding voltages to be read on the Pirani gauge. These two sets of measurements were carried out (Tables I and II). A short extrapolation to lower temperatures of the vapour pressure/temperature curve gives us simultaneous values of pressure and voltage function for the same temperature, and thus the necessary data for calculating α.

The method of obtaining the low temperatures was to immerse a small bulb containing a little pure carbon tetrachloride in an open-mouthed Dewar vessel containing various pure substances in the melting state. In such vessels, plugged with cotton wool, the rate of melting of these substances was slow and the carbon tetrachloride pressures remained constant over periods of 15 to 30 minutes. The substances were first solidified by means of liquid oxygen. Most of the refrigerating substances employed have melting points accurately determined by Timmermans, Van der Horst and Onnes[*] at Leyden and recommended by them for serving as temperature standards. In the following tables these substances are marked thus (L) and the melting points are accurate to within 0·1° C. Pressures are in millimetres of mercury.

---

[*] 'Comptes Rendus,' vol. 174, p. 365 (1922).

Table I.

Refrigerant.	M.P., ° C	Pressure of sat. vapour (direct readings).	1/T.	$\log_{10}$ P.
Carbon tetrachloride . .	−22·9 (L)	8·19	0·003998	0·9133
Mercury	−39·0	2·35	0·004274	0·3711
Chlorbenzene	−45·2 (L)	1·31	0·004390	0·1183
Ethyl malonate	−50·0	0·88	0·004480	$\bar{1}$·9450

Fig. 1 (A) shows the plot of $\log_{10}$ P against 1/T, from which the following relation between pressure and temperature was deduced :—

$$\log_{10} \mathrm{P} = 8 \cdot 982 - 2018/\mathrm{T}. \qquad \text{(A)}$$

Table II.

f (*v*). Voltage functions corresponding to sat. v.p. of carbon tetrachloride at low temperatures.

Refrigerant.	M.P , ° C.	*f* (*v*) (P. gauge readings).	1/T.	$\log_{10} f(v)$.
Chloroform	−63·5 (L)	5·930	0·004773	0·7751
Solid carbon dioxide	−78·0	1·159	0·005125	0·0641
Ethyl acetate	−83·6 (L)	0·563	0·005279	$\bar{1}$·7505
Toluene	−95·1 (L)	0·113	0·005620	$\bar{1}$·0531
Carbon disulphide	−111·6 (L)	0·010	0·006195	$\bar{2}$·0000
Ethyl alcohol	−117·0	0·0035	0·006410	$\bar{3}$·5441

Fig. 1 (B) shows the plot of $\log_{10} f(v)$ against 1/T, from which the following relation between voltage function and temperature was deduced :—

$$\log_{10} f(v) = 10 \cdot 308 - 2000/\mathrm{T}. \qquad \text{(B)}$$

The two sets of results plot as parallel straight lines. The parallelism of these lines is significant. Carbon tetrachloride occurs in two forms* with a transition point at − 48·5° C. corresponding to 1/T = 0·00445. From the parallelism of the lines A and B it is to be inferred that under the conditions of rapid cooling set up the transformation did not take place, as otherwise line B would have taken a position indicating lower vapour pressures than were found. The unstable form of carbon tetrachloride with the higher vapour pressure must therefore have persisted below the transition temperature.

---

* Latimer, 'J. Amer. Chem. Soc.,' vol. 44, p. 90 (1922).

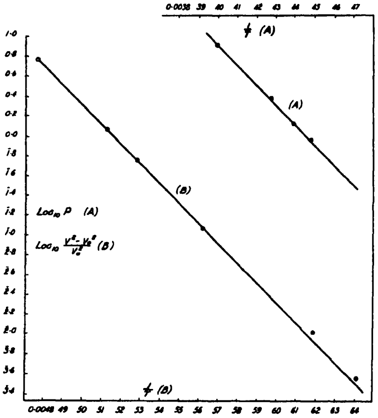


Fig. 1

The parallelism of A and B also shows the constancy of α over the pressure range covered.

Value of α.—From the equation of the gauge we see that

$$\log_{10} \alpha = \log_{10} P - \log_{10} f(v).$$

From equations A and B considered for the same temperature we have

$$\begin{aligned} \log_{10} P - \log_{10} f(v) &= 8\cdot982 - 10\cdot308 \\ &= \bar{2}\cdot6740, \end{aligned}$$

whence

$$\alpha = 0\cdot0472.$$

Remarks on the Use of the Gauge.—The filament temperature is about 250° C. and the stability of any substance under the conditions of measurement must be ascertained before the gauge is used for that substance.

The gauge zero varies slightly with room temperature, falling about 3 millivolts per degree rise in temperature. It is much more affected by jars to the apparatus, and these should therefore be avoided as far as possible. The maximum observed variation due to both causes was from 3·9 to 4·1 volts. In the technique about to be described it is necessary for these and other reasons to determine the gauge zero before each pressure measurement. This procedure and the avoidance of jars during measurements obviate any error due to change of gauge zero.

(3) *Apparatus and Technique.*

Fig. 2 shows the apparatus, which is of glass throughout, with the exception of the silica vessels C and K.

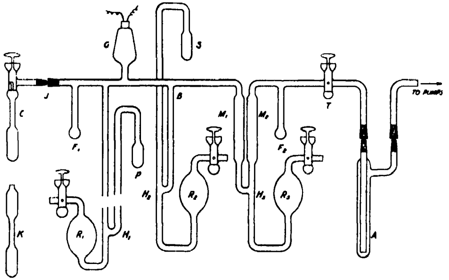


Fig. 2.

C is the charcoal container, made of silica and having a capacity of 12 c.c. It is furnished with an accurately ground end-on vacuum tap and a side tube ground to fit into the cup at J. The unlubricated joint is secured tight by means of Everitt's hard vacuum wax.

K is a sealed silica vessel, of approximately the same weight, volume and surface as C, and is used as a counterpoise in weighing the latter. The use of such a counterpoise ensured much more consistent weighings.

G is the Pirani gauge lamp.

F_1 and F_2 are small freezing tubes about 22 cm. in length.

P contains about 5 grams of coconut charcoal. This charcoal, when cooled in liquid oxygen, is used for the thorough clean-up of the apparatus and for the removal of gases displaced from the experimental charcoal.

The T-piece H_1 is about 90 cm. below the bend over of the tube leading to P. R_1H_1, R_2H_2 and R_3H_3 are mercury-seal devices for the T-pieces H_1, H_2 and H_3.

The bulb S contains about 4 c.c. of carbon tetrachloride. The top of the bulb is 90 cm. above the T-piece H_2, so that when the air is let back into the apparatus, S may still remain sealed off. To ensure this a pressure slightly in excess of atmospheric is adjusted in R_2 by means of a hand compression pump.

M_1 M_2 are close parallel tubes 2 cm. in diameter,* used as a U manometer in obtaining the pressure readings in Table I. M_2 was evacuated and the difference in height of the mercury columns was read on a cathetometer to within 0·05 mm.

A is a detachable liquid oxygen trap used for freezing out vapours and so protecting the pumps.

The tube from A was sealed to the evacuating system, which consisted of an electrically heated metal mercury condensation pump backed by a " Cenco " rotary oil pump. The apparatus could be reduced to $< 10^{-5}$ mm. of air in 30 minutes. The tap T served to withstand atmospheric pressure when the pumps were not in use.

The tap lubricant used was made by heating together vaseline, crepe rubber and paraffin wax in the ratio 16 : 10 : 1 by weight.

Except during evacuation the seal H_3 was kept closed so that the only lubricated tap in communication with the gauge was that on the charcoal container. This tap could not be dispensed with. The reduction of greased surfaces to a minimum by the substitution of mercury seals for taps greatly contributed to the success of the technique, for it was necessary that pressures due to foreign vapours should not exceed 10^{-4} mm,

The position of the mercury was adjusted by varying the pressure in the bulbs R_1, R_2 and R_3.

Temperature Control.—For the isotherms the container C was immersed nearly up to the cup in an electrically controlled thermostat the temperature of which was maintained at 25·0° C. ± 0·05°. For the isosteres the container was immersed in a Dewar vessel containing water at the required temperature. The natural rate of cooling for a vessel of this kind is small, and this rate was at each temperature just balanced by means of a small electrically heated coil. The mouth of the vessel was plugged with cotton wool. For 0° C. crushed ice and water were used. Temperatures could easily be adjusted at the required value and maintained constant to within 0·1° C. in this way.

---

* Large diameter to minimise capillary errors.

Pre-Evacuation of the Charcoals.—The weight of the clean silica container was noted; the charcoal was introduced followed by a known weight of tap lubricant. After connecting up C with a separate vacuum line the bulk of the air and moisture was removed. The charcoal was then raised to the desired temperature in a tubular electric furnace, evacuation being continued until the pressure fell to the pre-arranged value. Two evacuation temperatures were employed, some of the charcoals being evacuated at 110° C. and others at 800° C The 110° charcoals were evacuated to approximately 2×10^{-4} mm. and the 800° charcoals to 2×10^{-3} mm. measured at the evacuation temperature. These pressures were as low as could be reached in a reasonable time of evacuation. The standard pressure method ensures a more standardised state of the charcoals than does evacuation on a time standard. The furnace temperature was recorded either by means of a mercury thermometer or a thermo-couple, according to the temperature employed.

Zero Weight.—The weight of the clean container and the grease being known, the loss of weight due to the removal of the air from the container in evacuation was calculated. The excess over the weight of the greased, evacuated container gave the weight of evacuated charcoal, and the excess over the weight of the evacuated container and charcoal gave the amount of carbon tetrachloride sorbed. About 2 grams of charcoal were taken for experiment in all cases.

Purification of Carbon Tetrachloride.—Carbon tetrachloride, supplied as pure, was freed from traces of sulphur compounds by the method of Grüss,* the essential feature of which is the passage of chlorine through the boiling substance for several hours. The liquid was then washed with sodium carbonate solution, dried and redistilled. The dissolved air had to be eliminated very completely in view of the low-pressure measurements it was desired to make. This was carried out as follows: The carbon tetrachloride was solidified at − 183° C. in the bulb S and the air pumped off. S was then sealed off and the carbon tetrachloride melted and resolidified. The released air was again pumped off. This cycle was repeated eight to ten times until the air pressure over the liquid had fallen to 10^{-4} mm. It was found to be an advantage to solidify the carbon tetrachloride when high-pressure air was subsequently let back into the apparatus. The solution of any air that entered S was thus prevented and this air was then easily removed in one operation.

---

* 'Zeit. Elektrochem,' vol. 29, p. 144 (1923).

The Clean-Up Charcoal—It was essential that this charcoal should not develop a measurable pressure when at liquid oxygen temperature, as it was used for a final clean-up of the apparatus and also to sorb displaced gases during the pressure measurements. It was evacuated in the first instance at room temperature to 10^{-3} mm. gas pressure, the vapour content of the charcoal being unimportant. In this state it showed no gas or vapour pressure at liquid oxygen temperature. It was subsequently evacuated at room temperature to the above degree at frequent intervals.

Control of Spacing of Points.—In order to study the form of the isotherms it was necessary to have control over the position of experimental points. Considerable difficulty was at first experienced in the matter, mainly due to the large volumes of carbon tetrachloride vapour which charcoal sorbs at low pressures. (One gram of coconut charcoal at 25° C. sorbs about 950 litres at 2×10^{-3} mm.) In order to economise time, it was necessary to employ pressures somewhat higher than those at equilibrium, and these were adjusted as follows :—

(i) *For Sorption Points* (increasing concentration).—A standard charging pressure of $0 \cdot 3$ mm. was adopted which gave a convenient rate of charging. This was adjusted by cooling the bulb S by raising around it a vacuum bottle of liquid oxygen* until the gauge indicated the voltage corresponding to $0 \cdot 3$ mm. The carbon tetrachloride was then admitted at this pressure for the requisite period of time as indicated by a previous calibration.

(ii) *For Desorption Points* (decreasing concentration).—In this case the carbon tetrachloride pressure over the charcoal was increased by raising the temperature of the latter, which was then evacuated. The highest temperature employed was 100° C. (for the lowest concentrations). The desorption was carried on to a definite pressure at the elevated temperature which could be calculated by a simple formula from the pressure desired at 25° C. In this case one cannot work on a basis of time owing to the continual fall in the carbon tetrachloride concentration.

The spacing of points on the isotherms (Part II of this paper) will show the utility of these methods of control.

---

* Cooling by cold air only.

(4) *Effects due to Gas Content of Charcoal.*

It was observed that carbon tetrachloride displaced gases from charcoal even after a very thorough evacuation at 800° C. and re-evacuation at 25° C. to zero pressure. Pure carbon tetrachloride was found to have no measurable vapour pressure at − 183° C. (liquid oxygen temperature), whereas charges of vapour taken from the charcoal would not completely condense at this temperature. In addition, hysteretic effects were met with which could not be explained by the presence of permanent gases and this led to the discovery of carbon dioxide in the displaced gases. This was identified by (*a*) being completely condensed to zero pressure at − 183° C.; (*b*) the pressure being unaffected at − 126° C.; (*c*) being completely absorbed by soda-lime, the behaviour in all respects being identical with that of pure carbon dioxide from sodium bicarbonate under similar conditions.

The inclusion of these gas pressures in the total pressure readings would have given rise to considerable error, as the gas and carbon tetrachloride pressures were often of the same order*; but a more serious error, caused by the presence of these gases in the charcoal, results through their greatly retarding the equilibrium between charcoal and carbon tetrachloride, giving rise to irreversible effects.† Reversible results could only be obtained after displacing these gases, and this could only be effected by employing a sufficient pressure of carbon tetrachloride and simultaneously raising the temperature to 90° to 100° C.‡ The charcoal was then evacuated for a short time and cooled. After this procedure, if sufficient carbon tetrachloride had been employed, the gas pressures fell to a much lower value, equilibrium between charcoal and carbon tetrachloride was rapidly attained, and the isotherm or isostere was reversible. Small residual gas pressures always persisted, however, even after the major displacement had been effected, and, being too large to be neglected, were measured and allowed for by the method given below.

The experimental evidence which led to the above conclusions and procedure will be fully presented with the results in Part II of this paper.

---

* *Cf.* Coolidge, *loc. cit.*

† McGavack and Patrick, 'J. Amer. Chem. Soc.,' vol. 42, p. 946 (1920).

‡ Harned, 'J. Amer. Chem. Soc.,' vol. 42, p. 372 (1920).

(5) *Measurement of Gas Pressures in the Presence of Carbon Tetrachloride Vapour.*

The principle of the method employed for differentiating between gas and vapour pressures is due to Campbell (*loc. cit.*) and depends on the isolation of the gas by condensing the vapour out of the mixture to zero pressure by means of a refrigerant. A small side tube from the apparatus is immersed in a suitable cooling bath and the whole of the vapour condenses into this tube at zero vapour pressure, while the gas is not condensed; as the side tube represents only a small fraction of the total volume, the contraction of the gas due to cooling can be ignored and its pressure can now be measured alone.

In the present case it was necessary to measure pressures due to carbon tetrachloride, carbon dioxide and permanent gas, and suitable refrigerants had to be found for the two former substances. The bulbed tube F_1 in fig. 2, of about 3 c.c. capacity, was used as the freezer, and it was found that carbon tetrachloride had no measurable vapour pressure at − 183° C. (liquid oxygen temperature). Pure carbon dioxide, obtained from sodium bicarbonate, was also found to have no vapour pressure at − 183° C. and thus the pressure of permanent gas could be measured by freezing out carbon tetrachloride and carbon dioxide into F_1 by means of liquid oxygen.* It was established that by now removing the gas by evacuation no loss of carbon tetrachloride or of carbon dioxide from the freezer occurred. This removal of gas was effected by exposing the coconut charcoal in P at −183° C. while the carbon tetrachloride and carbon dioxide were held at zero pressure in the freezer F_1. By sealing off the coconut charcoal and warming up the freezer to room temperature, the total pressure due to carbon tetrachloride and carbon dioxide could now be measured. It was further established that carbon tetrachloride had no measurable vapour pressure at − 126° C. (the melting point of methyl cyclohexane), while carbon dioxide did not tend to condense at this temperature below pressures of about 2 mm. (the saturation v.p. of carbon dioxide at − 126° C.) and consequently behaved as a permanent gas at pressures measurable on the Pirani gauge. The use of methyl cyclohexane as a refrigerant thus afforded a means of differentiating between low pressures of carbon tetrachloride and of carbon dioxide, it being only necessary after the removal of permanent gas to freeze out the carbon tetrachloride with this refrigerant and measure the residual pressure due to carbon dioxide. The pressure due to carbon tetrachloride was obtained by difference. The pressure found on revaporising the carbon tetrachloride after removing carbon dioxide was always much too high owing to the accumulation of extraneous vapours during the cooling of $F._1$

---

* The charcoal container is naturally closed before the refrigerants are used.

An example of an actual measurement will make the procedure clear. In what follows,

$f(v)$ = voltage function.
V = voltage reading.
V_0 = gauge zero.

Analysis of Charge of Vapour from Charcoal.

1. Freezer at room temperature; $V = 4 \cdot 58$; $f(v) = 0 \cdot 331$, total charge.
2. Freezer at $-183°$ C.; $V = 3 \cdot 985$; $f(v) = 0 \cdot 007$, permanent gas.

Gas removed by exposing the coconut charcoal at $-183°$ C. for five minutes.

3. Freezer at $- 183°$ C.; $V_0 = 3 \cdot 97, f(v) = 0$ (gauge zero).

The carbon dioxide is now vaporised by replacing liquid oxygen on freezer by melting methyl cyclohexane.

4. Freezer at $-126°$ C.; $V = 4 \cdot 09$; $f(v) = 0 \cdot 06$, carbon dioxide.
5. Freezer at room temperature; $V = 4 \cdot 57$; $f(v) = 0 \cdot 324$, CCl_4, and CO_2, from 4 and 5 we have :—

$$f(v) \text{ carbon tetrachloride} = 0 \cdot 264.$$

Since $\alpha = 0 \cdot 0472$, $P = 1 \cdot 25 \times 10^{-2}$ mm.

The value of the constant α is $0 \cdot 0320$ both for carbon dioxide and for the permanent gas* (CO ?), thus making their pressure in the preceding example to be :—

$$P \, . \, CO_2 = 1 \cdot 9 \times 10^{-3} \text{ mm.}$$
$$P \, . \, CO = 2 \cdot 0 \times 10^{-4} \text{ mm.}$$

It is not necessary to know these pressures for making the correction, but it is of interest to know how the gas pressures vary during the course of an experiment.

It should be remarked that the procedure here adopted for obtaining the $f(v)$ carbon tetrachloride by difference is not theoretically correct since the voltage function of a mixture is only equal to the sum of the voltage functions of each constituent measured alone if the α values are the same. In the case of carbon tetrachloride and carbon dioxide this is not so, but the error committed by assuming it to be so does not exceed the sensitivity of the gauge when the proportion of carbon dioxide is small. With much higher proportions of carbon dioxide than were found in the present work, a special calculation would have to be made.

---

* *Cf.* Campbell, *loc. cit.*

(6) *The Charcoals Investigated.*

The following charcoals were selected for experiment :—

A.—Air activated birchwood charcoal (British origin).

B.—$ZnCl_2$ activated pinewood charcoal (German origin).

C.—Steam activated coconut charcoal (American origin).

D.—Steam activated mixed nut charcoal (British origin).

E.—Steam activated mixed charcoal, chiefly from briquetted coal (British origin).

F.—Steam activated coconut charcoal (British origin).

For further details of charcoals A, B and C, see Hand and Shiels.*

These charcoals were granulated and the granules used for experiment were those which were passed by a sieve of 10 meshes to the inch, but retained by one of 12 meshes to the inch. The packing density of the granules was as follows :—A = 0·205, B = 0·298, C = 0·490, D = 0·587, E = 0·412, F = 0·490.

(7) *Experimental Procedure.*

After attaching the container at J the apparatus is evacuated and the extraneous vapours condensed for one hour into F_2 (fig. 2) and A, both cooled in liquid oxygen. In this time the air pressure will have fallen to $< 10^{-4}$ mm. and the vapour pressure to 10^{-3} mm., and the charcoal container tap may now be opened and the carbon tetrachloride concentration adjusted. If a sorption point is required the seal H_3 is closed and a charge of vapour admitted to the charcoal. If a desorption point is required the charcoal is heated with hot water and carbon tetrachloride condensed out into F_2 and A to the desired pressure value. After the adjustment of the concentration, the container and seal H_3 are closed and the gauge side of the apparatus is cleaned up for about 18 hours by means of the coconut charcoal in P, cooled in liquid oxygen. P is then sealed off and the experimental charcoal is put in communication with the gauge by opening the container tap for 10 minutes. The carbon tetrachloride and gas pressures are then measured in the manner described, after which the container is removed and weighed. Data for a point on the isotherm are thus obtained. For the isosteres, a set of pressure measurements, each for a different temperature, is obtained for every concentration value.

---

* 'J. Phys. Chem.,' vol. 32, p. 441 (1928).

In the course of the isostere measurements a minute change in concentration occurs through the removal of several charges of vapour; owing to the low pressure of the vapour, however, this change in concentration is small enough to be neglected.

Remarks on Gauge Zero and Clean-up.—The clean-up of the apparatus with coconut charcoal at liquid oxygen temperature produces a high vacuum with respect to gases, but traces of vapour persist. The extraneous vapour pressure did not exceed 10^{-4} mm. and for pressure measurements above 10^{-2} mm. was neglected. For the measurements below 10^{-2} mm. the extraneous vapour pressure was further reduced as follows:—(*a*) Some of the pressure was due to carbon tetrachloride vapour exuding from the seal H_2, which could be entirely obviated by solidifying the carbon tetrachloride supply. Accordingly, during measurements, the carbon tetrachloride was maintained solid by immersing the supply bulb S in liquid oxygen. (*b*) A further source of extraneous vapour was found to be the wax at the joint at J, and this appeared to give off water vapour. The pressure of this could be considerably reduced by cooling the bulb F_1 in solid carbon dioxide and ether, which was without effect on carbon tetrachloride pressures below 10^{-2} mm. (the pressure of saturated carbon tetrachloride vapour at $-78°$ C. being $5 \cdot 47 \times 10^{-2}$ mm.). The foregoing precautions resulted in a depression of the gauge zero by 10 millivolts compared with that reached after the clean-up of the apparatus by the cooled coconut charcoal. Cooling the bulb F_1 in liquid oxygen gave a gauge reading some 2 millivolts lower still, and this being the lowest that could in any circumstances be obtained was taken as the gauge zero. As stated, this was checked during every pressure measurement.

This investigation forms part of a more comprehensive one undertaken by Prof. A. J. Allmand, to whom the author desires to express his indebtedness for close interest and help.

Summary.

1. This paper consists of a detailed description of the apparatus and experimental technique by means of which the low-pressure sorption isothermals of pure carbon tetrachloride vapour on charcoal, in absence of foreign gases, were determined.

2. The pressures were measured by means of a Pirani hot-wire gauge, of which the calibration and method of use are fully described. The quantities sorbed were determined by direct weighing.

3. Various difficulties were encountered, due to the presence of traces of

foreign gases and vapours in the charcoals and in the apparatus. The elimination of these difficulties is fully discussed.

4. Six charcoals of varying origin were employed. The pressure limits worked between were $2\cdot3 \times 10^{-1}$ and 1×10^{-4} mm. of mercury. Most work was done at 25° C., but several series of isosteres were also obtained, which permitted of the indirect determination of isothermals at higher temperatures.

5. The paper is illustrated by two diagrams. The results obtained, and their discussion, will be communicated later.

The Structure of Topaz $[\mathrm{Al}\,(\mathrm{F, OH})]_2\,\mathrm{SiO}_4$.

By N. A. Alston, Samuel Bright Scholar, Manchester University, and J. West, John Harling Fellow, Manchester University.

(Communicated by W L. Bragg, F.R.S.—Received July 30, 1928)

1. In a recent paper on the structure of certain silicates,* Prof. W. L. Bragg and one of the authors discussed in considerable detail the rôle which the oxygen atoms—due to their relatively large size—played in these compounds. The structures examined fell roughly into two classes: one in which the O atoms were throughout packed as closely together as possible, the other in which close packing was local only. Many examples of structures which had actually been worked out were given, and it was suggested that topaz might possibly prove a further example of a close-packed arrangement of large ions.

An attempt to solve the structure of topaz on these lines proved unexpectedly difficult, and a detailed general quantitative examination was therefore carried out. A full account of this examination, which has now been completed, will be published later in the 'Zeitschrift fur Kristallographie.' We propose here to give but a general outline of the structure, desiring rather to draw attention to an interesting new type or variant of close packing which the study has revealed, and which suggests important considerations in the elucidation of complex structures still awaiting analysis.

---

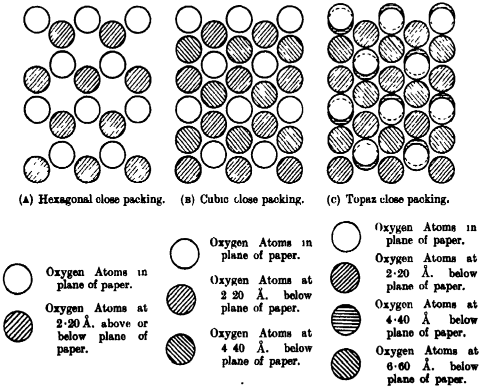
* 'Roy. Soc. Proc.,' A, vol. 114, p. 450 (1927). Topaz $[\mathrm{Al}\,(\mathrm{F, OH})]_2\,\mathrm{SiO}_4$ contains oxygen and fluorine atoms, but since the accepted sizes for the ions are roughly the same, we have, when considering the dimensional relations, regarded the two sets as identical, and referred to them as "oxygen" atoms.

2. *Forms of Close Packing.*

Two arrangements of spherically symmetrical atoms in closest packing are recognised—the so-called *cubic* and *hexagonal* forms. Although they are well known, it is convenient for our purpose to recall here certain of their characteristic features when the spheres are oxygen atoms of diameter 2·7 Å.

Both the cubic and hexagonal arrangements are such that they constitute an ordered collection of groups of four atoms arranged tetrahedrally and groups of six atoms arranged in octahedral fashion. It is within certain of these groups that are to be found in actual crystals the much smaller atoms such as Al and Si, and it is the distribution of these smaller atoms, which, by determining the three dimensional pattern characteristic of the crystal, controls the shape and size of the unit cell

The general relations between the two forms are probably best exemplified in figs. 1 (A) and (B), which represent views of the oxygen arrangement along the hexagonal axis of the hexagonal form and along the trigonal axis of the cubic form respectively.



(A) Hexagonal close packing. (B) Cubic close packing. (C) Topaz close packing.

○ Oxygen Atoms in plane of paper.
▨ Oxygen Atoms at 2·20 Å. above or below plane of paper.
○ Oxygen Atoms in plane of paper.
▨ Oxygen Atoms at 2 20 Å. below plane of paper.
▧ Oxygen Atoms at 4 40 Å. below plane of paper.
○ Oxygen Atoms in plane of paper.
▨ Oxygen Atoms at 2·20 Å. below plane of paper.
▤ Oxygen Atoms at 4·40 Å below plane of paper.
▧ Oxygen Atoms at 6·60 Å. below plane of paper.

FIG. 1.—Some Types of Close Packing.

The individual layers of atoms parallel to the plane of the paper are seen in both cases to possess an identical arrangement. The difference between the two forms lies in the way in which the successive layers are placed under any chosen layer. Thus, in fig. 1 (A) the relative dispositions of the planes of atoms are such that every *third* layer is exactly under the first layer (we are supposed to be looking in a direction perpendicular to the layers), whereas in fig. 1 (B) it is the *fourth* layer which comes exactly under the first. Certain important dimensional relations and directions are conveniently given in fig. 2.

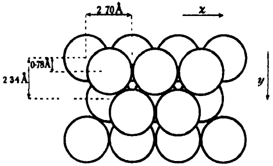


FIG. 2.—The x axis is perpendicular to plane of paper. The distance between layers of atoms parallel to plane of paper = 2·21 Å.

The three directions x, y and z are mutually perpendicular, the x axis being perpendicular to the plane of the paper.

Now, whatever the size and shape of the unit cell in an actual crystal may be, if the structure belong to either of the two close-packed types commonly recognised, it will possess the following characteristics :—

(*a*) The volume per oxygen atom will be of the order of 14 $Å^3$.

(*b*) The refractive index of the crystal will not be less than about 1·7.

(*c*) It will be possible to derive the unit cell by joining appropriate points in one of the two close-packed arrangements.

(*d*) A study of the X-ray spectra will reveal the presence of a certain group of outstanding spectra characteristic of, and corresponding to, certain simple spacings in one of the forms of close packing, although the intensities may be modified to some extent by the influence of the particular spatial distribution of metal and Si atoms.*

In order to illustrate some of these characteristics and to show the relations with topaz, we give in Tables I and II data (taken from the paper referred to above) referring to certain crystals whose structures—excepting topaz—

---

* See also Sir W. H. Bragg, ' Nature,' vol. 121, p. 327 (1928).

are all based on the hexagonal or cubic close-packed arrangement of oxygen atoms.

	Type of close packing.	Molecular volume M/ρ.	Volume per oxygen atom.	Refractive index (Na).
Ideal close-packed assemblage	—	—	13·94 Å.^{3}	1·71
Beryllium oxide, BeO	Hexagonal	8·26	13·62	1·726
Corundum, Al_2O_3	Hexagonal	25·6	14·05	1·768
Chrysoberyl, $BeAl_2O_4$	Hexagonal	34·4	14·18	1·749
Cyanite, Al_2SiO_5	Cubic	45·6	15·05	1·720
Spinel, $MgAl_2O_4$.	Cubic	39·5	16·30	1·724
Olivine, $(Mg, Fe)_2SiO_4$	Hexagonal	44·1	18·20	1·65 (Mg_2SiO_4)
Chondrodite, $H_2Mg_5Si_2O_{10}$	Hexagonal	100·2	18·10	1·62
Monticellite, $MgCaSiO_4$	Hexagonal	51·5	21·22	1·66
Topaz $[Al(F, OH)]_2SiO_4$	—	54·4	14·25*	1·61

* Fluorine and oxygen are regarded as equivalent.

In Table I we note that the molecular volume is in certain cases greater than 14 Å.^{3} (which represents a lower limit for oxygen atoms of diameter 2·7 Å.). This is to be ascribed to the distortional expansion caused by atoms such as magnesium, iron and calcium when introduced in the interstices of the oxygen framework, for which they are somewhat too large. This expansion alone will, of course, cause a reduction in the refractive index from 1·71 (see column 5), but the presence of the heavier ions causing the expansion may wholly or in part neutralise this effect. In the case of topaz the small value 1·61 is to be ascribed to the presence of fluorine atoms which, with oxygen atoms, form the framework of the structure.

Table II.

	x.	y.	z.
Hexagonal close-packed assemblage (orthorhombic cell)	4·42 Å.	4·68 Å.	2·7 Å
$BeAl_2O_4$.	$a = 4·42$	$b = 9·39 = 2 \times 4·70$	$c = 5·47 = 2 \times 2·74$
$(Mg, Fe)_2SiO_4$	$a = 4·76$	$b = 10·21 = 2 \times 5·10$	$c = 5·89 = 2 \times 2·94$
$MgCaSiO_4$	$a = 4·82$	$b = 11·08 = 2 \times 5·54$	$c = 6·37 = 2 \times 3·18$
$[Al (F, OH)_2SiO_4$—			
Case (i) . ..	$a = 4·64$	$b = 8·78 = 2 \times 4·39$	$\left.\right\} c = 8·38 = 3 \times 2·79$
Case (ii)	$b = 8·78 = 2 \times 4·39$	$a = 4·64$	

In Table II we give the dimensions of the unit cells of certain orthorhombic crystals which exhibit hexagonal close packing. In order to bring out clearly the way in which these cells fit into the arrangement of fig. 1 (A), we choose from fig. 1 (A) a unit orthorhombic cell, containing four atoms, whose axes (see fig. 2) are $x = 4\cdot 42$ Å $y = 4\cdot 68$ Å, $z = 2\cdot 7$ Å. These axes are given in Table II, where their relations to the sides a, b and c of the unit cells of the crystals will be evident. In the case of topaz, if we assume true hexagonal close packing, symmetry considerations force us to identify the a, b and c axes of the crystal with the x, y and z directions (case (i) in Table II), although dimensionally the agreement is more satisfactory if we suppose the b and a axes to be respectively parallel to the x and y directions (case (ii)). We shall see later that case (ii) and not case (i) is correct.

3. *Topaz.*

Topaz $[Al(F, OH)]_2 SiO_4$ is orthorhombic. The unit cell, which contains four molecules, has sides of lengths $a = 4\cdot 64\,(1)$ Å., $b = 8\cdot 78\,(3)$ Å., $c = 8\cdot 37\,(8)$ Å.

Space Group.—Unfortunately, crystallographers are not agreed as to whether the symmetry is holohedral or polar. Generally speaking, whilst the geometrical data suggest holohedral symmetry, the existence of pyroelectric and piezoelectric effects seem to point to polar symmetry. These latter effects are, however, admittedly very weak when present, and the direction of the suggested polar axis varies widely in different specimens.*

If the crystal class be orthorhombic bipyramidal, the X-ray data show the space group to be V_h^{16}.† This space group requires reflexion planes parallel to our c face and also possesses centres of symmetry. A structure based on this space group requires 15 parameters to define it. The alternative space group, if the symmetry be polar, is C_{2v}^9, and may be derived from V_h^{16} by removing the reflexion planes and centres of symmetry. A structure based on this space group is defined by 27 parameters.

In view of the character of the crystallographic data we adopted V_h^{16}, although during the actual analysis of the structure we have been alive to the possibility of the true space group being C_{2v}^9; our final results have justified this procedure.

---

* The crystal used in the present investigation was a topaz from Nigeria. We may, however, remark that no differences were noticed between our rotation photographs and those obtained by Leonhardt for Mexican topaz.

† Niggli, 'Geometrische Kristallographie des Diskontinuums,' p. 201.

An Apparent Possibility.—At first, general considerations appeared to point to the structure being based on the hexagonal close-packed arrangement of " oxygen " atoms. Criteria (*a*) and (*b*) above are satisfied, since the volume per " oxygen " atom is 14·2 Å.^{3}[*] and the refractive index is 1·61 (as already pointed out this low value of 1·61 may be ascribed to the presence of fluorine). It was not possible to fit the cubic form of packing into the cell under the given circumstances, whereas one position was possible for the hexagonal form, viz., that in which the *a*, *b* and *c* axes of topaz were respectively parallel to the *x*, *y* and *z* directions of fig. 2 (see Table II, case (i)). Thus criterion (*c*) was satisfied. There remained criterion (*d*). The results here were interesting. Of the expected " close-packing " spectra only those parallel to the *c* face were obtained ; the structure explained quite easily and in detail the whole of the (*o o l*) spectra. On the other hand, it was in direct conflict with certain outstanding spectra from other crystal planes and, in particular, failed to explain the striking (*h o h*) series (see Table III). It was therefore obvious that whilst the true structure undoubtedly possessed many features normally associated with closest packing, it also differed from the two types, described above, in some fundamental manner.

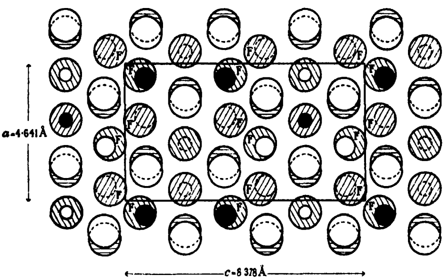
The Actual Structure.—The structure ultimately found as a result of detailed quantitative analysis is illustrated in fig. 3, which represents a view along the *b* axis. It possesses holohedral symmetry, and whilst it applies strictly to the particular crystal specimen examined, we believe that it also represents essentially the structure for all specimens of topaz. The oxygen and fluorine atoms combine to form a close-packed arrangement which belongs wholly neither to the cubic nor to the hexagonal type ; it is discussed below. The Al and Si atoms are distributed in a diffuse manner through the structure, each Al atom being surrounded by six " oxygen " atoms and each Si atom by four.

It is not easy in a 15-parameter structure to distinguish between two such atoms as oxygen and fluorine, but our results appear to indicate definitely that the four atoms about the Si atoms are oxygen, whilst of the six atoms round each Al atom four are oxygen and two are fluorine.

The " Oxygen " Arrangement.—In order to bring out more clearly the interesting way in which the oxygen and fluorine atoms are arranged, we give in fig. 1 (c) an idealised representation of the " oxygen " arrangement—viewed along the *b* axis—in which, of the slight distortions given in fig. 3, only those are retained which assist in making visible a layer of atoms which would otherwise be hidden. The figure should be compared with figs. 1 (A) and 1 (B).

---

[*] Leonhardt, ' Z. Krist.,' vol. 59, p. 216 (1923).



○ Oxygen Atoms at 4·40 A. above plane of paper.

▨ Oxygen Atoms at 2·20 Å. above plane of paper.

▤ Oxygen Atoms in plane of paper.

▧ Oxygen Atoms at 2·20 Å. below plane of paper.

● Aluminium Atoms at 3·24 A. above plane of paper

○ Aluminium Atoms at 1·15 Å. above plane of paper.

● Silicon Atoms at 3·90 Å. above plane of paper.

○ Silicon Atoms at 0·49 Å. below plane of paper.

Fig. 3.—Structure of Topaz viewed along *b* axis. The rectangle indicates the projection of the unit cell on the *b* face.

It will be seen that we again get a series of layers identical with each other and with those in figs. 1 (A) and (B), but the way in which the successive layers are disposed relative to a chosen layer differs from that in both (A) and (B). Beginning with the top layer in the plane of the paper, the first, second and third layers imitate fig. 1 (A). The fourth layer, however, departs from the arrangement to be expected from the sequence in fig. 1 (A), causing the second, third and fourth layers to resemble rather fig. 1 (B). The fifth layer falls under the first, and the series begins again. The packing in topaz may thus be regarded as a combination of the hexagonal and cubic forms. We may also show the simplicity of the arrangement by deriving it from the hexagonal form alone. Examining fig. 1 (C) we note that the first, second and third layers constitute a block identical (except for the slight distortions introduced to differentiate superimposed layers) with fig. 1 (A). But so also do the third, fourth and fifth layers. These two blocks are simply related in the structure, one being the reflexion of the other in a glide plane parallel to the xz plane (the x, y and z directions are identical with those given in fig. 2). The glide is parallel to the x direction, its length being equal to the depth of each block (4·4 Å.) (see also Table II, case (ii)). Had the two blocks been combined to form the normal hexagonal arrangement throughout, we should have had to place it in the unit cell—in order to satisfy the requirements of V_h^{16}—in such a way that the a and b axes of topaz would have been interchanged in their relations to the x and y directions (Table II, case (i)).

5. In providing an example of a type of closest packing which belongs strictly to neither of the two normally recognised types, the present structure suggests the possible existence of further types. These may all be regarded as examples of the various ways in which the single layers of closely packed

Table III.—Comparison of Experimental Results with those calculated from the Structure adopted.

Indices.		F* deduced from experiment.	F calculated from structure.
002	*c* face	2·6	3·5
004		5·2	5·5
006		31·1	36·2
008		1·4	2·7
0010		3·2	4·4
0012		8·3	11·2
0014		2·5	0·7
0016		2·9	2·5
0018		4·1	4·6
101	(101) face	0	0·3
202		0	0·3
303		35·2	34·4
404		0	0·2
505		0	1·6
606		7·7	10·1
707		0	0·3
808		0	1·6
909		0	1·4

---

[*] F denotes the ratio of the amplitude of the wave scattered by one molecule of topaz [Al (F, OH)]$_2$SiO$_4$, situated in the unit cell of the crystal, to that scattered by a single electron under the same conditions.

atoms in figs. 1 (A), (B) and (C) may be closely stacked along what we have called the x direction to form a periodic succession. The cubic and hexagonal types may then more properly be regarded as the two simplest cases of such stacking.

The effect of the possible existence of these further types is to give greater flexibility in the application of criterion (*c*) above—and therefore, of course, of criterion (*d*). We may expect these more complex types to be present in the structures of lower symmetry; for example, it is quite possible that certain of the complex triclinic silicates are based on one or other of these arrangements. It certainly seems advisable—since the analysis of a crystal structure is enormously simplified when the arrangement of the oxygen framework is known—to examine criterion (*c*) for a series of possible types of close packing before proceeding with the general analysis.

In conclusion, we wish to thank Prof W. L. Bragg, F.R.S., for his continued interest and advice during the course of the work.

NOTE.—After the present paper had been prepared for the press, the authors had the privilege of seeing an advance copy of a paper shortly to be published by G. Pauling and J. H. Sturdivant on the structure of brookite (TiO_2). It is interesting to find that the oxygen arrangement in this crystal—allowing for considerable distortion—is similar to that found in topaz.

Summary.

A quantitative analysis by X-rays of the structure of topaz $[\mathrm{Al}(\mathrm{F},\,\mathrm{OH})]_2\mathrm{SiO}_4$ has been carried out. Although this crystal is sometimes considered to belong to the orthorhombic pyramidal (polar) class, the structure actually found is holohedral in character.

The chief feature of the structure is the arrangement of the oxygen and fluorine atoms. Regarding these atoms as equal in size, they form a close-packed assemblage which belongs strictly to neither of the two well-known hexagonal and cubic types of close-packing. These two types and the assemblage found in topaz may be conveniently regarded as the simplest examples of the ways in which a series of identical planes, consisting of similar atoms in contact, may be closely stacked together, one on top of the other, so as to form a series of layers in periodic succession.

It is suggested that some of the more complex structures still awaiting analysis, which, whilst exhibiting certain features characteristic of close

packing, belong neither to the hexagonal nor to the cubic type, may actually prove to be based on one of these less simple arrangements.

Although in a structure of this kind it is difficult to distinguish between oxygen and fluorine atoms, it is believed that the four atoms which surround tetrahedrally each silicon atom are oxygen, whilst of the six atoms arranged symmetrically about each aluminium atom four are oxygen and two are fluorine.

The Rate of Emission of Alpha Particles from Radium.

By H. J. J. Braddick, B.A., Coutts Trotter Student of Trinity College, Cambridge, and H. M. Cave, M.A., Exhibition of 1851 Scholar of Queen's University, Kingston.

(Communicated by Sir Ernest Rutherford, P.R.S.—Received, July 31, 1928.)

1. *Introduction.*

A knowledge of the number, Z, of α particles disintegrations taking place in unit mass of radium in unit time is of considerable importance in the interpretation of radioactive changes, and, in particular, of the energy relations involved.

The heat evolution of radium and its short-lived decomposition products has been studied by a number of workers. Most of the heat production is accounted for by the energy of the α particles and recoil atoms, and in any particular experiment an allowance may be made for the β and γ ray energy absorbed. The experimental results are in agreement with the energy calculated from the number and energy of the α particles if a value is assumed for Z of about $3 \cdot 7 \,.\, 10^{10}$. If, on the other hand, the value $3 \cdot 57 \,.\, 10^{10}$ obtained by Rutherford and Geiger, or the value $3 \cdot 40 \,.\, 10^{10}$ recently published by Geiger and Werner, be taken, the calculation leaves a considerable portion of the heating effect unaccounted for, and this would involve an unidentified heat-producing mechanism in the disintegration.

2. *Measurements of Other Workers.*

The methods available for the determination of the number Z may be enumerated as follows :—

1. Direct counts.
   - (*a*) Electrical counters.
   - (*b*) Scintillation methods.
   - (*c*) Photographic methods.
   - (*d*) Wilson cloud chamber.
2. Indirect methods.
   - (*a*) Measurement of the total charge carried by the α rays, together with a knowledge of the charge on a single α particle.
   - (*b*) Measurement of the volume of emanation in equilibrium with radium.
   - (*c*) The rate of production of helium from radium.
   - (*d*) The rate of growth of radium from ionium.

A direct determination was made by Rutherford and Geiger* using an electrical counter in which the ionization produced by single α particles was magnified by collision. The active deposit of radium was used as a source, and the value obtained was $3 \cdot 57 . 10^{10}$ (on the International Standard).† The number of α particles counted was about 900, so that the probable error, excluding systematic effects, was about 3 per cent.

Hess and Lawson‡ made a determination by the electrical counter method, using active deposit sources. They obtained the value $3 \cdot 72 . 10^{10}$ but their work is susceptible of certain criticisms. Considerable changes in the number counted were produced by small changes in the experimental conditions. The counting was performed visually and there was a considerable systematic difference between the numbers obtained by the two observers.

Geiger and Werner§ made a careful determination by the method of scintillation counting with a zinc sulphide screen. They tested a large number of the individual crystals for sensitivity before applying them to the screen. They obtained the value $3 \cdot 40 . 10^{10}$ by counting 30,000 particles. Their work has been criticized on several grounds.‖ It is claimed that the experiments were vitiated by the occlusion of emanation in the platinum capillary tubes of the emanation sources used, that the efficiency of the screens was probably reduced by the addition of binding materials, and that the method of "double counting" used gave rise to systematic errors.

---

* Rutherford and Geiger, 'Roy. Soc. Proc.,' vol. 81, p. 141 (1912).

† Rutherford, 'Phil. Mag.,' vol. 28, p. 320 (1914).

‡ 'Wien. Akad. Ber.,' vol. 127, p. 405 (1918).

§ 'Z. Physik,' vol. 21, p. 187 (1924).

‖ Hess and Lawson, 'Z. Physik,' vol. 24, p. 402 (1924), and 'Phil. Mag.,' vol. 148 p. 200 (1924).

Geiger and Werner also made a determination with an electrical counter.* They showed that the electrical counter recorded all the α particles falling upon it and obtained the value $3 \cdot 48 \,.\, 10^{10}$ which they regard as an upper limit for Z.

Of the indirect methods, the measurement of the total charge carried by the α rays is the only one which can be used for an accurate determination.. Rutherford† made a measurement of the charge in 1905. Rutherford and Geiger‡ made an accurate measurement of the charge carried by the α particles from active deposit sources, and the result when corrected to the International Radium Standard and reduced with the aid of Millikan's value of the electronic charge gives $3 \cdot 48 \,.\, 10^{10}$. In those experiments a magnetic field was applied to suppress the effects of β rays and δ rays, and the experiment was performed in a fairly good vacuum. There was a considerable change in the observed α-ray current when an electrostatic field was imposed between the collecting electrode and the diaphragm near it. This was ascribed to ionization of residual gas and the mean of the currents observed with opposite electrostatic fields was taken as the true α-ray current. This procedure seems open to some question in view of the results of our preliminary experiments (see below).

Danysz and Duane§ made rather similar experiments, using thin-walled emanation bulbs as sources. Their value of the charge, reduced by the use of the Millikan number gives $3 \cdot 17 \,.\, 10^{10}$. This low value is almost certainly due to stopping of some of the α particles by their sources.

Since our work was begun, Jedrzejowski‖ has published what appears to be a careful determination of the charge carried by the α rays. The sources used were of radium B and C prepared by a method of dissolving Ra active deposit in acid and evaporation on a small glass plate. The charge was measured by a quartz piezo-electrique, and the γ-ray activity of the source was observed continuously during a run. Residual effects due to β rays were estimated by bringing a thin foil, sufficient to stop the α rays, into the path of the latter. The result obtained for Z was $3\cdot50 . 10^{10}$.

---

* Tätigkeitsber, 'Phys. Tech. Reichsanstalt,' vol. 47 (1924), and 'Verh. D. Phys. Ges.,' vol. 12, p. 12 (1924).

† 'Phil. Mag.,' vol. 10, p. 193 (1905).

‡ 'Roy. Soc. Proc.,' vol. 81, p. 162 (1908).

§ Amer. J. Sci., vol. 35, p. 295 (1913).

‖ 'Ann. Physique,' vol. 9, p. 128 (1928).

At the suggestion of Sir Ernest Rutherford we undertook a determination by the "total charge" method, and our result is $Z = 3\cdot68 . 10^{10}$.

3. *Experimental.*

In the determination of the number of α particles by the "total charge" method it is necessary to secure—

(1) A suitable radioactive source which may be compared with a radium standard.

(2) A system for securing a known fraction of the total α-ray emission.

(3) A system for collecting and measuring the charge.

Active Sources.—(*a*) The source used must be comparable with a radium standard by γ-ray methods.

(*b*) All the active material must be capable of firing α particles over the whole area of the limiting diaphragm without interference by other parts of the source or by the thickness of the active layer.

(*c*) The distribution of the active material must be geometrically definite so that the fraction of the α particles which pass the limiting diaphragm may be calculated.

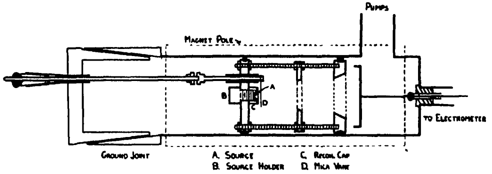
In our work sources of radium active deposit were used. The Radium A was allowed to decay before measurements of the α-ray activity were begun, and the α rays were therefore due entirely to Radium C. The γ rays used to compare the source with the standard came also almost entirely from Radium C, as is discussed below.

In order to obtain active deposit sources without possibility of contamination on the back, radium emanation was confined over mercury in a small glass tube with a hemispherical end round which a scratch had been made before introducing the emanation. After the tube had been left for some hours, the emanation was pumped off and the tube detached from the apparatus. It was washed with alcohol and baked for a few minutes in a vacuum furnace at a temperature of about 400° C. The hemispherical end was then cracked round with a hot wire, waxed to a brass carrier and broken off from the rest of the tube.

Several sources prepared by this method were tested for emanation contamination by sealing them up in glass tubes immediately after cracking off. They were then left overnight and tested in the morning with a γ-ray electroscope. In no case was any trace of emanation found, though 1/5000 of the initial activity could readily have been detected.

In order to calculate the effective mean solid angle subtended by the diaphragm at the source, the surface of the latter was divided into zones of 0·2 mm. depth. The distance of the centre of each zone from the diaphragm was calculated, and the solid angles obtained for the zones were each weighted with the area of the zone concerned, since the zones near the cracked edge were in general, incomplete. It was assumed that the surface was uniformly coated with active deposit, but the effect of departure from this condition would be small, since the difference in solid angle between the extreme zones of a source was about 7 per cent. The correction necessary to the solid angle for points on the source lying off the axis of the system was evaluated by the method of zonal harmonics.* It was about 1 part in 300 for a point lying on the extreme edge of one of our sources, and about 1 in 650 for a source as a whole.

The Measurement of the Charge.—The arrangement of apparatus for the charge collecting experiments is indicated in the figure. The source, attached to a threaded brass plug, was carried on a brass framework, the "source system," which carried also the diaphragm for defining the beam of α particles and which slipped inside the outer evacuated cylinder. The diaphragm was a brass disk with a carefully machined hole, bevelled as shown to reduce the scattering of α particles at the edge of the hole. A second perforated disk lying between the source and the diaphragm had a hole too large to obstruct the beam, but it served to stop particles scattered at the walls of the chamber, to stop β rays, (in conjunction with a magnetic field) and to strengthen the structure mechanically.



Two such source systems were constructed with diaphragms of different sizes. In one case the diameter of the diaphragm hole was about 2 cm., the distance from source to diaphragm about 4·5 cm., and the calculated solid angle about 0·012 of a complete sphere. In the other case the diaphragm diameter was reduced to 1·5 cm. and the solid angle to about 0·006 of a complete sphere. The dimensions of the diaphragm systems were measured several times during the work by means of a comparator.

---

* "Maxwell, Electricity and Magnetism," Chap. XIV.

The source system carried also an arrangement by which a mica vane could be turned over the source without disturbing the vacuum, so as to stop the α particles and allow the detection of any other effect producing a charge on the collecting electrode.

Over the source was screwed a little cap covered with thin aluminium leaf which served to prevent the contamination of the chamber by recoil atoms.

The front diaphragm was usually covered with thin metal leaf and the whole system was insulated from the chamber so that an electrostatic field might be applied between the diaphragm and the collecting electrode to test for ionization currents. The source system was placed in a position inside the chamber defined by a pin which entered a slot in one of the insulators.

A number of collecting electrodes were tried. The first was a flat disk of thin copper foil, but it was shown that δ rays escaped from it and falsified the result. Subsequent collectors took the form of shallow boxes made of copper foil, and a table of the results obtained with them is given below.

The charge collected was measured by the Townsend compensation method. The electrode was connected to the electrometer by a fine wire passing down the centre of a glass tube with an earthed silver coating inside. This tube was connected with the high vacuum of the chamber so that ionization inside it was negligible. The electrometer was a Compton instrument used at a sensitivity of about 3,000 divisions per volt. It was used as a null instrument, the charge communicated to it by the α particles being balanced by the charge induced on a condenser connected to a potentiometer. This arrangement greatly reduces electrical leakage from the insulated system and enables a sensitive electrometer to be used for the accurate measurement of quite large charges.

The charge induced is given by

$$q_{AB} : V$$

where V is the difference of potentials applied to the condenser at the beginning and end of a reading, and q_{AB} is the coefficient of capacity between the insulated system and the variable potential side of the condenser. The other capacities associated with the system are not involved. Our condenser, an air condenser, was enclosed in an evacuated bell jar to avoid ionization leaks. It was compared on several occasions *in situ* with a standard condenser lent by the N.P.L., using the electrometer itself as an indicator. The voltage applied to it in an experiment was measured by a Weston voltmeter, which was calibrated at several points against a standard cell.

The chamber containing the source and collecting electrode was evacuated by a Gaede steel two-stage diffusion pump suitably backed and provided with liquid air traps. The vacuum was estimated by the discharge in a large discharge tube connected to a 6-inch coil. This tube gave a faint green fluorescence in all the experiments, the vapour of tap grease being always present.

The chamber was placed between the poles of an electromagnet which gave, with 6 amperes current, a field of about 5,000 gauss.

The γ-ray Measurements.—The γ-ray measurements were made with an ionization chamber covered with 2 cm. of lead. It was set up near the experimental chamber and connected to an electrometer and compensating condenser by a wire passing down a long evacuated tube. Readings were made in the same way as for the α-ray current, the voltage applied to the compensating condenser being adjusted by a set of calibrated resistances so that the time for a reading was between 35 and 100 seconds. All readings were reduced by the use of factors derived from the resistance calibrations. These were made on a Tinsley potentiometer and were repeated several times during the work. By this arrangement sources of very different activity could be compared with some accuracy. Tests made by adding the effects of two sources showed that the ionization obtained was proportional to the γ-ray activity to an accuracy of about 1 part in 700. The natural leak of the ionization chamber was about $1 \cdot 5$ per cent. of the activity due to the standard source. It was measured by observing the rate of deflection of the electrometer and measuring the sensitivity of the latter.

Preliminary Experiments.—Preliminary experiments were carried out, using a thin-walled emanation tube as a source. With a flat collecting electrode it was found that considerable changes in the apparent α-ray current (order of 5 per cent.) were produced by applying electrostatic fields between the diaphragm and the collecting electrode, and that a small distortion of the latter gave rise to a large change in the α-particle current. These effects were finally traced to δ rays which left the foil approximately along the lines of the magnetic field, and the flat electrode was replaced by a shallow copper cup.

There was now a very much smaller effect on applying an electrostatic field. Forty volts had no measurable effect; -200 volts applied to the diaphragm produced a change of $0 \cdot 6$ per cent. in the apparent α-ray current and $+200$

volts, a change of 2·8 per cent. This was attributed to small ionization currents in residual vapours. When the vacuum was made poorer by stopping the diffusion pump the effect of the fields on the apparent α-ray current increased, while retaining its asymmetry with respect to direction of field. Within considerable limits, however, this change in vacuum had no effect on the magnitude of the α-ray current without applied electrostatic field, and we therefore assume that it has no effect on our results.

It was found that the observed α-ray current was independent of the magnetic field within the limits of about 3,500 and 6,000 gauss. With the magnetic field on and the α rays cut off by a mica vane just thick enough to stop them, the residual effect was a slow drift in the same direction as the α-ray current. This was probably due mainly to the expulsion of electrons from the collector by γ rays. It was observed in the experiments with Radium C sources and subtracted from the observed α-ray current as a correction.

It was found that the thin foils used to cover the front diaphragm (copper leaf—calculated stopping power 1·3 mm. air; or aluminium leaf, 0·8 mm.) had no effect on the observed α-ray current.

Experimental Procedure with an Active Deposit Source.—In an experiment the γ-ray electrometer was first calibrated by a standard source placed in a standard position defined by lines in the focal plane of a small microscope. The source prepared in the way described above was placed in this standard position, usually about 25 minutes after removal from the emanation, the time of removal being taken as a zero of time for the experiment. Readings were taken on the γ-ray electrometer for about 20 minutes, the time from zero, as defined above, to the end of each reading being recorded. At the end of this time the Radium A remaining on the source was negligible and the source was transferred as quickly as possible to the experimental chamber. Readings were resumed on the γ-ray electrometer while the chamber was being pumped out and continued throughout the run, so that a decay curve was obtained for the Radium C while it was actually in the chamber.

As soon as the discharge tube showed the vacuum to be good enough, α-ray readings were begun. At the beginning and end of each observation the stop watch was pressed when the electrometer spot was passing steadily through zero, so that the small lag of the electrometer had no effect on the readings. Between these times the electrometer was kept within a few millimetres of zero by manipulating the potential divider. Several times during the experiment the mica vane was brought over the source and the movement of the electrometer measured with the α particles cut off. The effect observed was a slow drift in the same direction as the α-particle current. It was usually about 1 per cent. of the latter, and was subtracted in estimating the α-particle current. A certain number of α-ray readings were taken with potential differences of + 40 and − 40 volts applied to the source system. No certain differences in the apparent α-ray current were found.

At the end of the experiment the source was again transferred to the standard position and its decay followed for a further 15 minutes. Meanwhile the activity transferred to the little cap by recoil was measured by a γ-ray electroscope—it was usually about 1 per cent. of the activity of the source on removal from the chamber, and no correction was necessary for the difference in solid angle subtended by the diaphragm at the cap and at the face of the source. A correction was applied in fitting the decay curves as described below. Finally the standardisation of the γ-ray electrometer was repeated—the two standardisations usually agreeing to about 1 part in 800.

From these observations curves were plotted showing the α-ray activity and the γ-ray activity during the middle part of the experiment, and the γ-ray activity of the source measured under standard conditions at the beginning and end of the experiment. By an extrapolation across the short time gaps corresponding to the transference of the source from the standard position to the chamber and *vice versa*, the ratio of the γ-ray currents produced by the same source in each of these two positions was found, and hence from the γ-ray readings taken within the chamber it was possible to deduce the activity of the source expressed in terms of milligrams of radium at any time during the experiment. In determining the ratio at the lower end of the curve a correction was made for the active matter transferred from the source to the recoil cap. The two ratios usually agreed to about 1 part in 500.

It was found that the decay curve of the active deposit did not in general follow the calculated curve, based on the work of Bracelin,* for Radium A, B, C, in equilibrium. The equilibrium was apparently disturbed by the conditions of exposure. The published decay was, however, used as a basis on which to construct an extrapolation curve.

The ratio $\frac{\alpha\text{-ray activity}}{\gamma\text{-ray activity}}$ should, of course, remain constant during the experiment. When calculated from the curves it showed, in general, unsystematic deviations of a few parts in 1,000 and a mean value was taken for the calculation of the result.

---

* 'Proc. Camb. Phil. Soc.,' vol. 23, p. 150 (1926).

Numerical Example of an Experiment.—February 6, 1928. Calibration of γ-ray electrometer with radium standard—times for reading—

Before run 101·0, 101·0, 100·8, 100·7, 101·1, 101·1, 101·1, 100·9, 101·0, 101·0 seconds.

After run, 100·9, 100·7, 100·9, 101·0, 100·9, 100·8, 101·1, 101·0, 100·8, 100·9 seconds.

Mean of all—$100 \cdot 9_5$.

Voltage across potential divider from resistance box calibration, 0·3339.

γ-ray activity in arbitrary units (millivolts per second),* 3·308.

Natural leak, 0·048.

Corrected γ-ray activity, 3·260.

Simultaneous readings of α-ray and γ-ray activity. (Former single readings in "millivolts per second" applied to the α-ray compensating condenser; latter taken from a smooth curve and expressed in millivolts per second as in the case of the standard.)

Time.	α.	γ.	α/γ.
56·17	2·504	2·541	0·985
59·28	2·367	2·396	0·988
62·16	2·251	2·273	0·990
66·46	2·142	2·170	0·987
73·00	1·826	1·853	0·985
76·05	1·726	1·738	0·993
79·53	1·587	1·612	0·985
83·43	1·477	1·487	0·993
		Mean	0·988

Ratio $\dfrac{\gamma\text{-ray activity in standard position}}{\gamma\text{-ray activity inside chamber}}$ by extrapolation of curves

At beginning = 3·56.

At end = 3·55 (including correction for γ-ray activity transferred to recoil cap).

Solid angle calculated from geometrical dimensions as fraction of complete sphere = 0·01211.

Calculation of Result.

(Capacity of condenser in α-ray measuring system, from mean of calibrations, 897·6 cm. Value of standard in equivalent γ-ray activity = 6·38 mg.)

---

* Applied to the γ-ray measurement condenser.

Charge communicated by the α particles from 1 gram of radium in 1 second

$$= \frac{897\cdot6 \times 0\cdot988 \times 3\cdot260}{6\cdot38 \times 300 \times 3\cdot55_5 \times 0\cdot01211} = 35\cdot09 \text{ E.S.U.}$$

Assuming Millikan's value of e, charge on a single α particle

$$= 2 \times 4\cdot774\cdot10^{-10} \text{ E.S.U.}$$

Therefore the number of α particles per second from 1 gram of radium

$$= \frac{35\cdot09}{2 \times 4\cdot774\cdot10^{-10}} = 3\cdot67_5 \,.\, 10^{10}.$$

When small corrections as discussed below are taken into account the result becomes

$$3\cdot69_2 \,.\, 10^{10}.$$

Results.—The results of 16 experiments are given below, and the variations in the apparatus for different experiments indicated.

Date.	Collector.	Source System.	Result.
December 7, 1927	1	1	$3\cdot67_8 \,.\, 10^{10}$
December, 13, 1927	1	1	$3\cdot67_1$
December 14, 1927	1	1	$3\cdot67_4$
December 17, 1927	2	1	$3\cdot69_3$
January 17, 1928	2	1	$3\cdot68_8$
January 25, 1928	2	1	$3\cdot68_9$
January 26, 1928	2	1	$3\cdot70_5$
February 6, 1928	3	1	$3\cdot69_2$
February 9, 1928	3	2	$3\cdot69_4$
February 17, 1928	3	2	$3\cdot69_5$
February 22, 1928	3	2	$3\cdot70_3$
March 15, 1928	4	1	$3\cdot69_1$
March 22, 1928	4	1	$3\cdot63_3$
March 26, 1928	4	1	$3\cdot66_0$
March 27, 1928	4	1	$3\cdot67_8$
March 29, 1928	4	2	$3\cdot67_3$
	Mean	of all	$3\cdot68_2 \,.\, 10^{10}$

The first experiments were made with a collecting electrode of thin copper foil 0·022 mm. thick—calculated stopping power for α particles 9·3 cm. It was in the form of a shallow open cylinder of depth about 3 mm. (Collector 1). It was thought that with so thin a collecting electrode the collection of fast β rays and the photoelectric effect of the γ rays would be negligible. We found that these effects, as measured by the current with α-rays cut off, were small with all the collectors used. Collector 2 was of similar shape and was spun from copper sheet, 0·12 mm. thick. Collector 3 was made from the thin foil and was about 1 cm. deep. It was lined at the bottom with a second layer of the copper foil. In some experiments it was lined with platinum foil 0·017 mm. thick (Collector 4).

It will be seen that there is no certain difference between the results obtained with these different arrangements.

The solid angle was also changed in the experiments marked "Source System 2," in which the smaller diaphragm was used. The agreement of these experiments with the others shows that scattering of α rays at the edge of the diaphragm is negligible.

A small correction had to be applied for absorption of the γ rays in the glass sheath used to support the γ-ray standard during measurements. This amounted to 0·25 per cent. in the result. Using the results of Slater* a small correction was applied for the effect of the Radium B in the source in producing γ-ray ionization. The ionization produced through 20 mm. of lead by the Radium B in a source in equilibrium with radium is about 0·4 per cent. of the total. In our sources there was of course less Ra B than the equilibrium amount, and the necessary correction in comparing them with the standard was taken as $-$ 0·2 per cent. According to Henderson† about 1/170 of the particles which leave a bare source deposited on platinum are singly charged. Assuming that this is true for our sources, a correction of 0·3 per cent. was applied to our result. The stop-watch used for the α-ray readings was compared occasionally with a standard clock, and it was found to gain consistently by 0·19 per cent. The correction necessary was applied in our result.

The α-ray current was measured directly and it was found with a steady source (an α-ray tube) that successive readings might be relied on to agree within 1 part in 300, and it seems unlikely that the errors in the measurement of the α-ray current can affect the result of a single experiment to a greater extent than 0·3 per cent.

The γ-ray measurements are rather less certain. They are made in terms of the laboratory radium standard, which has recently been tested at the N.P.L. and certified as 6·38 mg. ± 1 per cent. The measurements involve the comparison of sources of widely differing activity, and we consider that the γ-ray measurements may lead to uncertainty of 0·5 per cent. in the result of a single experiment (excluding uncertainty in the standard).

We consider the error involved in the measurement of the solid angle to be less than 0·2 per cent.

---

[*] 'Phil. Mag.,' vol. 44, p. 300 (1922).

[†] 'Roy. Soc. Proc.,' A, vol. 109, p. 157 (1925).

4. *Discussion.*

The number Z, obtained in these experiments, lies between the value of Hess and Lawson on the one hand, and the ones found by most of the other workers on the other.

The results of the heating experiments of a number of workers are collected in Meyer and Schweidler's "Radioaktivität,"* and they have been compared with the heating calculated from the energies of the α, β and γ rays in a recent paper by Watson and Henderson.† The velocity of the α particles from Ra C was taken from Rutherford and Robinson‡ and Briggs§ and the velocity of the α particles from the other products obtained from the Geiger|| v^3 relation. The energy of the β rays was found by Gurney¶ and that for the γ rays of Ra B and Ra C from the direct heating measurements of Ellis and Wooster.**

It appears that the heating calculated using our value for $Z = 3 \cdot 68 \,.\, 10^{10}$ is about 2 per cent. below that observed by Meyer and Hess†† for the case of radium and its short-lived decay products. It is about 1 per cent. below the values obtained by Rutherford and Robinson‡‡ for radium emanation and products of Ra C. It is also 1 per cent. below the heating observed by Hess§§ for radium free from its products. In this case there is no correction necessary for β rays, and the correction for the γ rays absorbed is small.

The heating calculated for $Z = 3 \cdot 68 \,.\, 10^{10}$ agrees to within 1 part in 300 with that observed by Watson and Henderson—in separate experiments on radium emanation + products, on Ra B + C and on Ra C. These experiments were made using the same radium standard as was used in our work. The accuracy claimed was 1 per cent.

It appears therefore that there is no necessity to invoke an unrecognized heat-producing mechanism in radioactive disintegration, and this conclusion agrees with that drawn by Watson and Henderson from their observation that the ratio of calculated and observed heating was constant over a number of radioactive elements.

---

* "Radioactivität" (1927 Edn.), p. 225 *et seq.*

† 'Roy. Soc. Proc.,' vol. 118, p. 318 (1928).

‡ 'Phil. Mag.,' vol. 28, p. 552 (1914).

§ 'Roy. Soc. Proc.,' vol. 118, p. 549 (1928).

|| 'Roy. Soc. Proc.,' vol. 83, p. 505 (1910).

¶ 'Roy. Soc. Proc.,' vol. 109, p. 540 (1925).

** 'Phil. Mag.,' vol. 50, p. 521 (1925).

†† 'Wien. Akad. Ber.,' vol. 121, p. 603 (1912).

‡‡ 'Phil. Mag.,' vol. 25, p. 312 (1913).

§§ 'Wien. Akad. Ber.,' vol. 121, p. 1419 (1912).

Alternative calculations of the heating effect of the β and γ rays have been made by Meitner* and Thibaud.† The former assumes that all the β and γ rays are originally produced with the highest of the observed energies and that the energy that they give up before leaving the atom appears as heat. She finds that the heating effect agrees with that obtained by assuming $Z = 3\cdot 50 . 10^{10}$. This calculation is, however, in disagreement with the experimental results of Ellis and Wooster on the heating of the γ rays of Ra B and Ra C, and of the β rays from Ra E,‡ and it is probable that the assumptions involved are invalid. Thibaud makes the heating effect agree with $Z = 3\cdot 57 . 10^{10}$, but his calculation also seems open to considerable doubt.

We are, however, unable to account for the difference between our results and those recently obtained by Jedrzejowski, using much the same method.

5. *Summary.*

The number of α particles emitted by 1 gram of radium in 1 second has been determined by collecting the charge carried by the α particles from active deposit sources. The number obtained is

$$3\cdot 68 . 10^{10} \pm 1 \text{ per cent.}$$

in agreement with recent values of the heating effect.

We are very glad to express our thanks to Sir Ernest Rutherford for his interest and advice throughout the work. We desire to thank also Mr. G. R. Crowe for the preparation of the radioactive sources. The work has been made possible for one of us (H. J. J. B.) by grants from Trinity College, the London County Council and the Board of Scientific and Industrial Research.

---

* 'Naturwiss,' vol. 12, p. 1146 (1924).

† 'C. R.,' vol. 180, p. 1166 (1925).

‡ 'Roy. Soc. Proc.,' A, vol. 117, p. 109 (1928).

The Structure of the Band Spectrum of Helium.—V.

By Prof. W. E. Curtis, D.Sc., and A. Harvey, B.Sc., Armstrong College,
Newcastle-upon-Tyne.

(Communicated by T. H. Havelock, F.R.S.—Received July 31, 1928.)

1. *Introductory.*

The analysis of this spectrum has now reached a point at which it becomes possible to predict the positions of new bands, sometimes precisely and sometimes only approximately, with some degree of confidence. There are two distinct and non-combining families of electronic levels which are closely analogous to those of atomic helium,‡ and each of these is accompanied by one or more relatively weak vibrational levels. The main features of the rotation structure of any particular band can also be predicted from a knowledge of the electronic levels involved. It therefore seemed obvious that in continuing this work§ attention should first be directed to locating and analysing the predictable bands which might be expected to appear on the spectrograms available. These having been accounted for the analysis of the remainder, if any, of the spectrum should be simplified.

Work was accordingly begun on these lines, and a search made for the bands 4D → 2P of the orthohelium family and for 3D → 2P and 3S → 2P of the parhelium family. It had not proceeded far, however, when it became apparent that the spectrum contained a number of other bands, some of them quite strong, which did not fit into the existing scheme as regards either location or structure. Most of these had been analysed into branches in the early stages (1922) of the investigation of this spectrum, but publication of the results was postponed until some interpretation of them could be obtained. This was now found to be possible with the help of more recent developments, and as the results promised to be of considerable interest it was thought preferable to abandon the original plan and to concentrate on the new bands. The present communication deals with four of these, together with one member of the " ordinary " orthohelium family.

---

‡ Mulliken, 'Proc. Nat. Acad. Sci.,' vol. 12, p. 158 (1926).

§ The previous papers of this series will be referred to as I, II, III, IV, and have appeared in these 'Proceedings' as follows :—I, vol. 101, p. 38 (1922) ; II, vol. 103, p. 315 (1923) ; III, vol. 108, p. 513 (1925) ; IV, vol. 118, p. 157 (1928).

Spectrograms from two sources have been utilised :—

(1) In the region λλ 6440 — 5710, first order of 10 feet concave grating, Imperial College, South Kensington. Dispersion about 5·4 Å. per millimetre.

(2) In the region λλ 4515 — 4458, third order of 7 feet 8 inches concave grating, Military College of Science, Woolwich. Dispersion about 2·3 Å. per millimetre. These were taken by Dr. Jevons in the course of an investigation of the Zeeman effect in this spectrum.‡

The order of relative accuracy of the wave-numbers in one band is estimated at about 0·05 cm.$^{-1}$ for lines of intensity 2 and over. In the case of the weakest lines the error is liable to be somewhat larger.

2. *Notation and Tables.*

Confusion due to the diversity of notations employed is of frequent occurrence in band spectrum work, and is particularly liable to arise in this case on account of the existence of certain complications. It will be as well, therefore, to state as concisely as possible the notation which will be used here.

The two families of bands, the electron levels of which correspond to the ortho- and parhelium atomic levels will be designated by the symbols $_o\mathrm{He}_2$ and $_p\mathrm{He}_2$ respectively. Band lines will be numbered by the rotational quantum number (j) of the *final* state, the electronic levels being denoted by a letter with figure prefixed in the usual way. Thus for the R branch of the band due to the transition 3S → 2P we shall write $\mathrm{R}(j) = 3\mathrm{S}(j+1) - 2\mathrm{P}(j)$, and so on. The electronic contribution to the wave-number will usually be omitted, since it disappears in forming term differences. The j values (here integral) of the new mechanics will be used.

The rotational levels of the P and D states are double, and will be distinguished by the subscripts A and B. Alternate members only are observed in each case, as follows :—

S levels :	$1_B, 3_B, 5_B, \ldots\ldots$	
P ,,	$1_A, 3_A, 5_A, \ldots\ldots,$	$2_B, 4_B, 6_B, \ldots\ldots$
D ,,	$3_B, 5_B, \ldots\ldots,$	$2_A, 4_A, 6_A, \ldots\ldots$
X ,,		$2_A, 4_A, 6_A, \ldots\ldots$
Z ,,	$1_B, 3_B, 5_B, \ldots\ldots$	

The subscripts are assigned according to the original scheme of Mulliken (*loc. cit.*) whereby A → A and B → B transitions give P and R branches, whilst A → B and B → A give Q branches. It may be necessary, as Mulliken has

‡ Curtis and Jevons, 'Roy. Soc. Proc.,' A, vol. 120, p. 110 (1928).

suggested,‡ to revise this in the light of recent theoretical developments, but it is probably somewhat premature to do so at present, while the analysis of the spectrum is still in progress.

It is convenient to retain the designations P, P′, Q_1, Q_2, R, R′ used in previous papers, their significance being as follows :—

The final rotational levels of P, R′ and Q_1 have odd j values, and those of P′, R and Q_2 have even j values.

The letter P is used to denote both a branch and a level, but it is not expected that this will cause any confusion, particularly as the latter will usually be given its appropriate suffix. Thus for the P branch associated with the 3X → 2P transition we shall write

$$P(j) = 3X_A \;\; (j-1) - 2P_A \;\; (j).$$

The suffix B has of course nothing to do with the B in the expression $Bj\,(j+1)$ for the rotation term.

The first vibrational level associated with a particular electronic level is indicated by placing (1) after the symbol for the latter, *e.g.*, 2P (1). Otherwise the zero vibrational level is to be understood.

Particulars of the new bands are given in Table I. Blends are denoted by * and perturbed lines by † in both tables and text throughout. Wave-number differences marked ** are derived from two blends.

Table I.—Details of New Bands.

The Band 3D (1) → 2P (1) of He_2.

j.	P′ branch.			Q_1 branch.			R branch.		
	Int.	λ (air) I.A.	ν (vac.).	Int.	λ (air) I.A.	ν (vac.).	Int.	λ (air) I.A.	ν (vac.).
1									
2							2	5739·33	17418·8
3				1+	5753·54	17375·8*			
4	Mas	ked by P9 of 0 → 0	[17319·2]				Mas	ked by $Q_1$7 of 0 → 0	[17449·4]
5				1	52·74	378·2*			
6	0	5780·62	294·4				1—	18 78	481·4
7				1	52·08	380·2*			
8	Mas	ked by P′15 of 0 → 0	[17269·3]				1—	08·14	514·0
9				2	50 13	386 1*			
10	1	5796·10	248·2						

---

‡ 'Phys. Rev.,' vol. 30, p. 791 (1927).

Table I—(continued).

j.	P branch.			Q_2 branch.			R′ branch.		
1							2	5744·32	17403·7
2				1+	5753·54	17375·8*			
3	Masked by P′8 of 0 → 0		[17334·9]				2	34·46	433·6
4				1	52·84	377·9*			
5	2	5776·64	17306·3				Masked by $Q_2$13 of 0 → 0		[17462·9]
6				1	52·08	380·2*			
7	1	85·34	280·3				1	14·57	494·3
8				1+	51·09	383·2			
9	1	93·85	254·9				1—	04·88	524·0
10				2	50·13	386·1*			
11	1—	5802·26	229·9						
12									
13	1—	10·66	205 0						
15	1—	19 25	179 6						

The Band 3X → 2P of $_3He_2$.

j.	P branch.			Q branch.			R′ branch.		
	Int.	λ (air) I.A.	ν (vac.).	Int.	λ (air) I.A.	ν (vac.).	Int.	λ (air) I.A.	ν (vac.).
1							2	6235·23	16033·48
3	1+	6263·76	15960·45				5	28 32	051·27
5	2+	79·57	920 26		Not observed.		6	23·65	063·31
7	2+	97·61	874 65				4+	20·98	070·19
9	1	6317·48	824·73				5	20·02	072·67
11	7	39·07	770 82*				1	20 59	071·21
13	0	61·79	714 52				1	22·48	066·32
15	0	85·20	655 82				3	25·60	058·28
17							0	29·81	047·43

* Blend with R10 of 3S → 2P.

Table I—(continued).

The Band 3Z → 2P of $_o$He$_2$.

j.	P′ branch.			Q branch			R branch.		
	Int	λ (air) I.A.	ν (vac.)	Int	λ (air) I.A.	ν (vac).	Int	λ (air) I A	ν (vac.).
1				7	5959 83	16774 38			
2	1	5970·20	16745·23				8	5954 17	16790·31
3				9	60 87	746 16			
4	2	90·65	688·06				10	58 76	777 38
5				9	84 98	703 87			
6	1—	6016 16	617 30				9	66 43	755 81
7				10	6003 37	652 71			
8	5	44 98	538 08*				10	76 19	728 45
9				8	23 85	596 10			
10	1—	75 76	454 29				10	87 58	696 63
11				8	45 98	535 36			
12							10	6000 29	661 26
13				4	69 49	471 29			
14							9	14 14	622 88
16							7	29 06	581 75
18							5	44 98	538 08*
20							3	61 86	492 04
22							1 –	79 61	443 87

P of $_o$He$_2$

j.	P′ branch.			Q branch			R branch		
	Int.	λ (air) I A.	ν (vac.).	Int.	λ (air) I A	ν (vac.).	Int	λ (air) I.A	ν (vac)
1		Not observed.		1	4457 91	22425 74			
2							6	4458 43	22423 16
3				1 ?	67 22	379 04			
4							8	63 32	398 60
5				3	78 03	325·02			
6							8 ?	69 10	369 59
7				5	89 88	266 10*			
8							8	75 56	337 31
9				1	4502 23	205 03			
10							6	82 51	302 67
11				0	15 16	141 40			
12							5	89 88	266 10*
14							3	97 54	228 16
16							1	4504 54	193 64†
18							0	14 81	143·13

The intensities of Q3 and R6 are unreliable owing to the proximity of ghosts.

Table I—(continued).

The Band 3Z → 2P of $_{p}\mathrm{He}_2$.

j.	P′ branch.			Q branch.			R branch.		
	Int.	λ (air) I.A.	ν (vac.).	Int.	λ (air) I.A.	ν (vac.).	Int.	λ (air) I.A.	ν (vac.)
1				1	6315·38	15829·98			
2							1 1+	6309·81 10·29	15843·96† 842·75†
3				3 3	27·76 28·21	799·02† 797·90†			
4							5	16·34	827·59
5				3	45·61	754·57			
6							6	25·40	804·91
7				3	66·56	702·74			
8							6	36·38	777·52
9		Not observed.		1	80·49	646·40			
10							10	48·84	746·56*
11				1	6413·87	586·92			
12							4	62·34	713·14
13				0	39·44	525·02			
14							2	76·79	677·55
16							1	91·96	640·33
18							0	6407·48	602·40

* Blend with R8 of 3S (0) — 2P (0).

3. *The Band* 3D (1) → 2P (1) *of* $_{o}\mathrm{He}_2$ (*near* λ 5750).

This is a weak counterpart of 3D → 2P of $_{o}\mathrm{He}_2$ (λ 5730) and is displaced about 60 cm.$^{-1}$ in the less refrangible direction. It is clear from its intensity, location and structure that it is due to the transition between the first vibration states associated with the 3D and 2P levels. Similar vibrational bands have been described in IV, and by Weizel and Füchtbauer‡ in the case of S → P and P → S transitions. The present observations are rather unsatisfactory owing to the faintness of the lines, incomplete resolution of the Q branches and confusion with lines of the main band.

The interpretation assumed may be checked by forming the following differences :—

(1) $R(j-1) - P'(j+1) = 2P_B(j+1) - 2P_B(j-1)$ for the first vibrational state.

These should agree with the corresponding values obtained from the band 3S (1) → 2P (1), near λ 6416, described in IV. The results are

For j =	3	5	7	9
Band λ 5750	99·6*	155·0*	212·1*	265·8
,, λ 6416	99·0	155·3	211·0	266·1

‡ 'Z. Physik,' vol. 44, p. 431 (1927).

(2) $Q_1(j) - P'(j+1) = R'(j) - Q_2(j+1) = 2P_B(j+1) - 2P_A(j)$ for the first vibrational state. The results are

For $j =$	1	3	5	7	9
Band λ 5750	27·9	56·2	83·3	111·0	137·9
,, λ 6416	28·2	56·3	83·9	111·0	137·8

All the values for λ 5750 are dependent upon at least one blend, but the agreement confirms the final level as 2P (1), and we may proceed to evaluate the initial term differences as follows: –

$3D_A(j+1) - 3D_A(j-1) = R'(j) - P(j) = 156{\cdot}6,\ 214{\cdot}0,\ 269{\cdot}1$, for $j = 5, 7, 9$.

$3D_B(j+1) - 3D_B(j-1) = R(j) - P'(j) = 130{\cdot}2,\ 187{\cdot}0,\ 244{\cdot}7$, for $j = 4, 6, 8$.

These are not numerous or accurate enough to justify the application of the usual methods for the evaluation of B, but an approximate value may be arrived at by comparison with the corresponding results for the vibrationless 3D term. These average respectively 4·05 per cent. and 4·1 per cent higher than the above, and the B values derived from them are 7·49 and 7·63. We may therefore take the 3D (1) B values to be about 4 per cent. less than those for 3D, *i.e.*,

For $3D_A$, $B = 7{\cdot}19$, and for $3D_B$, $B = 7{\cdot}32$.

The values of r_0, the nuclear separation, are $1{\cdot}076$ and $1{\cdot}066 \times 10^{-8}$ cm respectively. As in all other cases investigated the molecule expands when vibration sets in, and by much the same amount (*cf.* $\Delta r_0 = +0{\cdot}021$ here and $+0{\cdot}015$, $0{\cdot}018$ and $0{\cdot}016$ for 3S, 4S and 2P respectively).

The position of the band origin, ν_0, is also best determined by comparison of corresponding lines of the two bands, and is estimated at about 17376, that of the main band being at 17437. The difference of 61 cm $^{-1}$ between these two values gives approximately the difference between the vibration frequencies of the 3D and 2P states. The latter has been calculated as 1744 (see IV, p. 164) and the former is therefore 1683 cm $^{-1}$. This does not agree well with the value of 1505 previously calculated from the rotation term differences, but the latter is probably at fault since, as will be shown later, there are systematic errors in the O — C residuals which indicate that the formula used is not exact.

4. *The Band* $3X \rightarrow 2P$ *of* $_p\mathrm{He}_2$ (*near* λ 6250).

In the preceding paper of this series an account was given of a new band belonging to the $_o\mathrm{He}_2$ family and consisting only of R′ and P branches. The final electronic level was found to be 2P and the initial was provisionally designated 3X. Since it gives R′P branches in combination with $2P_A$ rotation terms it must also be labelled A.

The band now under consideration is weaker, but closely similar in structure. It likewise consists of R′P branches but no Q, and the R′ branch reaches a head between R′ (9) and R′ (11) in both bands. There is no way of directly checking the interpretation, since the 2P term differences of ${}_p\mathrm{He}_2$ have not been isolated, but there is a great deal of indirect evidence which supports the view that this band is 3X → 2P of ${}_p\mathrm{He}_2$. The X and P term differences obtained on this assumption correspond very closely with those of ${}_o\mathrm{He}_2$, the electronic term values fit well in the general scheme, and the intensity relatively to the ${}_o\mathrm{He}_2$ band is just about what experience of this spectrum would lead one to expect.

The final term differences, $2P_A(j+1) - 2P_A(j-1) = R'(j-1) - P(j+1)$ are as follows, the corresponding ${}_o\mathrm{He}_2$ values being added for comparison.

$j =$	2	4	6	8	10	12	14
${}_p\mathrm{He}_2$	73·03	131·01	188·66	245·46	301·85*	356·69	410·50
${}_o\mathrm{He}_2$	73·20	131·69	189·59	246·76	303·31	358·62	412·73

The molecular constants were derived as in previous cases, taking as the term formula $F(j) = B(j-\rho)^2 - \beta(j-\rho)^4$. In this case a satisfactory fit was obtained by neglecting ρ, the least squares values for B and β being 7·300 and $5 \cdot 32 \times 10^{-4}$ respectively. These are very close to the corresponding ${}_o\mathrm{He}_2$ values (7·334 and $5 \cdot 20 \times 10^{-4}$) and give the approximate value of the vibration frequency $\omega_0 (= 2\sqrt{B^3}/\beta)$ as 1710 cm.$^{-1}$ (*cf.* 1744 ${}_o\mathrm{He}_2$).

The initial term differences, $3X_A(j+1) - 3X_A(j-1) = R'(j) - P(j)$ are as follows—

$j =$	3	5	7	9	11	13	15
${}_p\mathrm{He}_2$	90·82	143·05	195·54	247·94	300·39*	351·80	402·10
${}_o\mathrm{He}_2$	92·11	144·73	197·25	249·7	301·69	352·91	403·42

The least-squares solution gives $B = 6 \cdot 5829$, $\beta = 1 \cdot 18 \times 10^{-4}$ and $\rho = 0 \cdot 0548$. The values are rather similar to those obtained for 3X of ${}_o\mathrm{He}_2$ (see IV), β is again abnormally low in comparison with other levels, giving an improbably high value for the vibration frequency $\omega = 2\sqrt{B^3}/\beta$.

The O — C residuals resulting from the least-squares treatment are also of some interest. In the first place they are, although not large, distinctly systematic in character, so that we must conclude that they are due at least in part to inaccuracy of the formula itself. In this respect they are quite different from the P and S terms so far investigated, which show no trace of systematic discrepancies. Further, they are closely similar to the corresponding residuals for the $3X_A$ terms of ${}_o\mathrm{He}_2$, and also to those of the $3D_A$ and $3D_B$ terms of ${}_o\mathrm{He}_2$, provided that in the former case the signs of the residuals are reversed.

The actual values for the first six term differences are as follows :—

$3X_A$ of ${}_p\mathrm{He}_2$	+0·086	−0·075	−0·122	+0·055	+0·144	−0·087
$3X_A$ of ${}_o\mathrm{He}_2$	+0·145	−0·173	−0·107	+0·037	+0·195	−0·117
$3D_A$ of ${}_o\mathrm{He}_2$	−0·055	+0·069	+0·083	−0·079	−0·105	+0·076
$3D_B$ of ${}_o\mathrm{He}_2$	+0·063	−0·063	−0 073	+0·022	+0·116	−0·064

The similarity will be more apparent from the plotted values exhibited in fig. 1 The $3D_A$ values are plotted with reverse signs

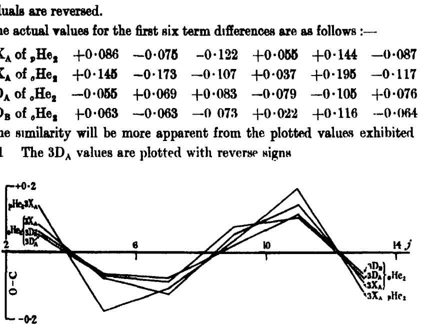


Fig. 1.—O — C Residuals for Term Differences.

It is clear that the formula used, which gave an entirely satisfactory representation in the case of P and S states (see III) requires a slight modification which is essentially of the same character for the 3X and $3D_B$ terms, and of a complementary nature for the $3D_A$. The deviations are not attributable to the omission of the σ from the full term formula

$$B\{(j-\rho)^2-\sigma^2\}-\beta\{(j-\rho)^2-\sigma^2\}^2$$

since the effect of this is only to produce a slight change in the value of B. It is more probably the onset of the effect which in the Z terms causes the above formula to break down completely for low j values, as will be shown in the next section

5. *The Z → P Bands.*

Three bands of this type have been identified up to the present, two belonging to the ${}_o\mathrm{He}_2$ family and the other to the ${}_p\mathrm{He}_2$. In view of their unusual appearance and characteristics it will be as well to give a general description of them before stating any detailed results They resemble one another so closely that it will suffice to describe 3Z → 2P of ${}_0\mathrm{He}_2$; this is not only the strongest but is also the least confused by overlapping of other bands.

Quite casual examination of the spectrum in the region about λ 6000 is sufficient for the recognition of two strong series both degraded towards the red. The intensity distribution in each case indicates that this is also the direction in which the rotational quantum number increases. But the wave-number intervals between the lines, instead of showing a linear variation with ordinal number, as in the case of all other known helium bands, decrease much more rapidly in the neighbourhood of the origin. The relative disposition of the two branches is again very different from any previously met with in this spectrum, so that it is not apparent whether they should be regarded as two P branches or two Q branches, or in some other way. The question was settled by the discovery of a combination principle relating these branches to the 2P level of ${}_0\mathrm{He}_2$. If we anticipate the conclusions reached below and designate the strong and weak branches R and Q_1, respectively, then it was found that $R(j) - Q_1(j+1) = 2P_A(j+1) - 2P_B(j)$ as derived from the differences $R(j) - Q_1(j+1) = Q_2(j) - P(j+1)$ of the band 3D → 2P of ${}_0\mathrm{He}_2$ (λ 5730). The actual values are (for $j = 2, 4, \ldots 12$)

$R(j) - Q_1(j+1)$ from this band =

44·15, 73·51, 103·10, 132·35, 161·27, 189·97.

$2P_A(j+1) - 2P_B(j)$ from λ 5730 =

44·04, 73·61, 103·06, 132·30, 161·16, 189·99.

This establishes 2P as one of the electronic states responsible for this band, and as it falls so close to 3X → 2P and not far from 3S → 2P it is safe to assume that 2P is the final state. This assumption is supported by the reasonable value (2·935) it gives for the initial effective quantum number. If this initial state is called 3Z, it follows from the combination result given above that $R(j)$ and $Q_1(j+1)$ must have the same initial term, which may be either $3Z_B(j+1)$ or $3Z_A(j)$. In the former case the observed branches will be R and Q_1, and in the latter Q_2 and P. Several considerations favour the latter view, for example :—

(1) It makes the Q branch the stronger, as in all other helium bands.

(2) Both branches are degraded to the red. If one of them is an R branch it must turn (*i.e.*, reach a head) in the neighbourhood of the first line.

(3) The first of the two branches has the smaller intervals, and thus looks more like a Q branch than the latter.

Definite evidence as to which is the correct view is obtainable by calculating a third branch (P′ or R respectively) from the two observed ones with the help of the 2P rotation differences derived from λ 5730. On doing this the first five lines of the calculated P′ branch were all found within 0·15 cm.$^{-1}$ of the calculated positions, but the alternative procedure (calculation of R branch) gave no lines within 0·55 cm.$^{-1}$ of observed ones, and only one within 1·05 cm.$^{-1}$. Four of the five P′ lines are weak and the abnormal strength of the other is accounted for by its coinciding with R (18). The combination differences obtained from the new branch are

$$Q(j) - P'(j+1) = 2P_B(j+1) - 2P_A(j),$$

the values being as follows for $j = 1, 3, 5, 7, 9$

From this band, 29·15, 58·10, 86·57, 114·63*, 141·81.
From λ 5730, 29·37, 58·01, 86·50, 114·48, 141·90.

We conclude, therefore, that this band consists of R, Q_1 and P′ branches. Other possible branches were looked for but none could be identified.

The strength of this band 3Z → 2P suggests that the next member 4Z → 2P of the series should be strong enough to observe, probably in the neighbourhood of λ 4460. This is a region rather rich in lines, but there is one very prominent series‡ near He λ 4472 which strongly resembles the R branch above, and application of the $2P_B(j+1) - 2P_B(j)$ differences to these gave six lines of an associated Q_1 branch. The differences $R(j) - Q_1(j+1)$ given below are in good agreement with the previous values.

$R(j) - Q_1(j+1) = 44{\cdot}12,\ 73{\cdot}58,\ 103{\cdot}49^*,\ 132{\cdot}28,\ 161{\cdot}27.$

We have not succeeded in identifying a P branch in this case, but by analogy with the previous band it is likely to be too weak to observe. There appears to be a large perturbation in R (16) of this band, its wave-number being about 10 cm.$^{-1}$ too high in relation to the expected value.

The next member of this series would fall near to λ 4000, but has not yet been identified. It is probable that only the R branch would show on our plates, so that no combination differences would be obtainable. But there is a band of a very similar character near λ 6300 which is undoubtedly due to the transition 3Z → 2P in $_p\mathrm{He}_2$. A direct proof by means of the combination principle is not possible, since the necessary 2P rotation term differences for $_p\mathrm{He}_2$ are not all known,§ but the indirect evidence is very strong. Thus, apart from the general resemblance of the two branches to the Q and R branches of λλ 5950 and 4460, the initial and final term differences closely parallel the 3Z and 2P of $_o\mathrm{He}_2$, and the effective electronic quantum number (2·952) is very close to that of 3Z of $_o\mathrm{He}_2$ (2·935). Like all the other $_p\mathrm{He}_2$ bands it is much fainter than its $_o\mathrm{He}_2$ counterpart, and it is therefore not surprising that we have as yet been unable to locate the next member of the series, particularly as this probably falls in a region of some complexity (near λ 4670).

---

‡ Marked ζ in 'Roy. Soc. Proc.,' A, vol. 89, p. 146 (1913), and also mentioned by McLennan, Smith and Lea, 'Roy. Soc. Proc.,' A, vol. 113, p. 183 (1926).

§ The $2P_A$ differences only are obtainable from λ 6250.

6. *The Rotation Term Differences for Z → P Type Bands.*

(*a*) 3Z → 2P of $_o\mathrm{He}_2$ (near λ 5950).

The combination result leads with the notation employed to the following expressions for the three branches, the electronic contribution and quantum number being omitted for convenience :—

$$\begin{aligned} \mathrm{P}'(j) &= \mathrm{Z_B}(j-1) - \mathrm{P_B}(j) \\ \mathrm{Q}\ (j) &= \mathrm{Z_B}(j) - \mathrm{P_A}(j) \\ \mathrm{R}\ (j) &= \mathrm{Z_B}(j+1) - \mathrm{P_B}(j). \end{aligned}$$

Hence $\mathrm{Z_B}(j+1) - \mathrm{Z_B}(j-1) = \mathrm{R}(j) - \mathrm{P}'(j) = \Delta_1 \mathrm{Z_B}(j)$ say, for brevity. The values for $j = 2, 4, \ldots 10$, are

$$\Delta_1 \mathrm{Z_B}(j) = 45{\cdot}08,\ 89{\cdot}32,\ 138{\cdot}51,\ 190{\cdot}37^*,\ 242{\cdot}34$$

Since the P′ branch is so weak and one member is a blend a better set of values can probably be obtained by utilising the Q_1 and R branches only, thus :

$$\Delta_1 \mathrm{Z_B}(j) = \mathrm{R}(j) - \mathrm{Q}_1(j-1) + \{\mathrm{P_B}(j) - \mathrm{P_A}(j-1)\}.$$

Two independent values for the term difference in brackets can be derived from the band 3D → 2P (λ 5730) and a third from 3S → 2P (λ 6400). They are respectively $\mathrm{R}'(j-1) - \mathrm{Q}_2(j)$, $\mathrm{Q}_1(j-1) - \mathrm{P}'(j)$, and $\mathrm{Q}_1(j-1) - \mathrm{P}'(j)$, the numerical results being

$j =$	2	4	6	8	10	12
	29·36	58·07	86·44	114·42	141·85	168·74
		57·98	86·52	114·47	141·88	168·93
	29·38	57·99	86·55	114·56	141·96	168·70
Means	29·37	58·01	86·50	114·48	141·90	168·79

By adding these to the observed differences $\mathrm{R}(j) - \mathrm{Q}_1(j-1)$ we get the following values for $\Delta_1 \mathrm{Z_B}(j)$

$$45{\cdot}30 \qquad 89{\cdot}23 \qquad 138{\cdot}44 \qquad 190{\cdot}22 \qquad 242{\cdot}43 \qquad 294{\cdot}69$$

These will be adopted as the final values.

(*b*) 4Z → 2P of ${}_o\mathrm{He}_2$ (near λ 4460).

As the P branch is missing in this case the second method described above has been followed, with the results

$$\Delta_1 Z_B(j) = 27{\cdot}08 \quad 77{\cdot}57 \quad 131{\cdot}07 \quad 185{\cdot}69^* \quad 239{\cdot}54 \quad 293{\cdot}49^*$$

There is some doubt here as to whether Q (1) is correctly identified. It might alternatively be under R (2), in which case the first difference would be 29·37 instead of 27·08. But the tabulated value is preferred since it leads to a smooth curve representing the intervals between corresponding 3Z and 4Z term differences.

(*c*) 3Z → 2P of ${}_p\mathrm{He}_2$ (near λ 6300).

In this case not only is the P branch missing but the $2P_B$ term differences are unknown. But with the aid of the $2P_A$ differences derived from 3X → 2P (λ 6250) both $2P_B$ and 3Z differences may be evaluated as follows :

For 3X → 2P we have

$$R'(j) = X_A(j+1) - P_A(j) \qquad P(j) = X_A(j-1) - P_A(j)$$

For 3Z → 2P we have

$$R(j) = Z_B(j+1) - P_B(j) \qquad Q_1(j+1) = Z_B(j+1) - P_A(j+1).$$

Hence $P_B(j+1) - P_B(j-1) =$

$$R'(j) - P(j+2) - R(j+1) + Q_1(j+2) + R(j-1) - Q_1(j).$$

This gives for $j = 3, 5, \ldots 11$,

$$\Delta P_B(j) = 102{\cdot}89,\ 159{\cdot}51,\ 216{\cdot}51\ \ 273{\cdot}33^{**},\ 328{\cdot}21^*.$$

We may now calculate the required initial term differences

$$\begin{aligned} Z_B(j+1) - Z_B(j-1) &= R(j) - R(j-2) + P_B(j) - P_B(j-2) \\ &= Q_1(j+1) - Q_1(j-1) + P_A(j+1) - P_A(j-1) \end{aligned}$$

The resulting values, for $j = 2, 4, \ldots .\ 12$, are

$$\Delta_1 Z_B(j) = \begin{cases} 42{\cdot}07 \\ 40{\cdot}95 \end{cases} \begin{cases} 86{\cdot}56 \quad 136{\cdot}83 \quad 189{\cdot}12 \quad 242{\cdot}37^{**} \quad 294{\cdot}79^{**} \\ 87{\cdot}68 \end{cases}$$

The double values for $Z_B(3) - Z_B(1)$ and $Z_B(5) - Z_B(3)$ are due to a perturbation in $Z_B(3)$ whereby it is resolved into two components about equal in intensity. This is responsible for the doubling of R (2) and Q_1 (3), which both originate in this level. The close agreement of the component separations (1·21 and 1·12 respectively) indicates that this is actually a perturbation and not a merely accidental proximity of a foreign line in each case. Similar instances have been noted in 4P → 2S and associated bands of $_0He_2$.[‡] The only suspicious circumstance here is the rather high intensity of the Q_1 (3) doublet. In the other cases (R2 here, Q (9) of 4P → 2S, etc.) the sum of the component intensities is of the same order as the expected intensity, but in this case each component is itself of about the expected intensity.

7. *The Rotation Terms of the Z Level.*

The usual method of deriving term values from term differences is based upon the use of a formula representing term values as a function of j, *e g.*, $F(j) = B(j-\rho)^2 - \beta(j-\rho)^4$, which gives a fairly accurate representation of all previously known bands in this spectrum. In some cases ρ is negligibly small. In the case of the Z terms such a formula is quite inapplicable, at least for the lowest terms, and the actual term values cannot therefore be obtained. The nature of the deviation is well shown by plotting the second differences Δ_2F of the term values against j. If a formula of the above type holds this should give a parabola, and when β and ρ are small, as always in this spectrum, the axis approximately coincides with that of Δ_2F and the vertex is at $j = 0$, $\Delta_2F = 8B$. This gives in fact a rapid approximate method of estimating the value of B and thence the moment of inertia of the molecule. The curves obtained in this way for various electronic states are shown in figs. 2 to 5, the values of 8B derived from least-squares treatment of the data, where available, being plotted on the $j = 0$ axis.

For the S and P states the curves closely approximate to parabolas and extrapolate accurately to the calculated value of 8B. In the case of the 3D state the curves have a steeper slope and do not extrapolate well (especially $3D_x$) to 8B. The 3X curves depart still farther from the expected form, showing a tendency to reach a maximum somewhere around $j = 6$ or 7, and definitely refusing to extrapolate to the calculated 8B. The conclusion reached in Section 4, that the formula used is incapable of accurately representing the D and X rotation terms, is thus confirmed.

The 3Z and 4Z curves depart very widely from the others, rising steeply with increasing j and apparently tending to a constant value. It is quite possible, however, that if they could be followed to high enough j values they would tend to fall off like all the rest. But the curve differs so radically from the previous ones that it is quite impossible to represent it by a formula of the usual type. Deviations of a somewhat similar character are to be found in the case of some doublet and triplet bands (*e.g.*, OH, CH, NH), and have received a theoretical explanation ;‡ but this seems to apply only to terms showing electronic multiplicity, which these do not. The form of the deviation suggests

---

[‡] See III and 'Roy. Soc. Proc.,' A, vol. 120, p 110 (1928).

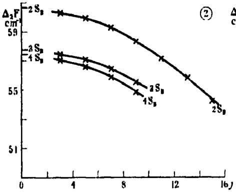


FIG 2

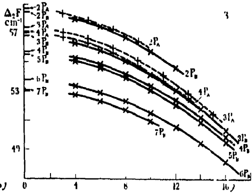


FIG 3

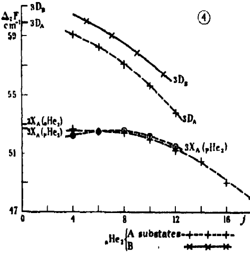


FIG 4

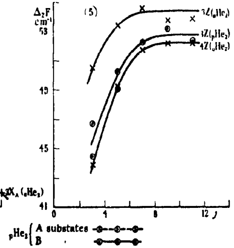


FIG. 5

that negative powers of j may occur in the expression for the rotation term. If this is the case the formula for the second differences will be so complicated that there is little hope of discovering an accurate empirical representation.

In the absence of a theoretical expression for the rotation term values we may retain the usual formula, and for the purpose of this discussion formally account for the deviation by supposing B to assume abnormally small values for small j values. The very unusual appearance of the branches may then be explained as follows :—

---

‡ Hund, 'Z. Physik,' vol. 36, p. 657 (1926); Kemble, 'Phys. Rev.,' vol 30, p. 387 (1927).

If the initial and final rotation terms are given by $B'j(j+1)$ and $B''j(j+1)$ it is easily shown that the R branch will turn at a value of j given by $(B'+B'')/2(B''-B')-1$. Thus for a given mean value of B the turning point will be earlier the greater the difference $B''-B'$. This is usually comparatively small, but cases are known in which the " head " is quite near to the origin (*e.g.*, near $j=2$ for β system of NO, near $j=3$ for iodine absorption bands). Here the " head " is in the neighbourhood of the first observed line ($j=2$), with the consequence that the R branch takes on the appearance of a Q branch. The unusually open appearance of the Q branch near the origin is readily accounted for in a similar way.

Another feature of these bands which calls for comment is the distribution of intensity between the branches. The R is much stronger than the Q, which is again much stronger than the P′. This is very different from the other helium bands, and, in fact, appears quite exceptional In all bands of this spectrum due to SP combinations the Q branch is distinctly stronger than either R or P. Further, in S → P transitions the P branch is stronger than the R, whilst in P → S the reverse is the case. In the D → P transition we have $P > Q > R$, and in the case of X → P the P and R are about equal, the Q being absent or very weak

The third anomalous feature of these bands is the nature of the Zeeman effect shown by them. This has only actually been observed for one, 4Z → 2P of ${}_o\mathrm{He}_2$, but may safely be presumed to be similar for the other two. A detailed description of the effects has been given elsewhere (*loc. cit.*) so that we need only mention here that the anomalous characteristics are—

(1) That the effects are larger than for other He_2 bands studied, and vary with the rotational quantum number in a different manner, and

(2) That the Q and R branches appear to show similar polarisation properties, unlike other bands.

8 *Electronic Levels in* He_2.

The number of electronic levels identified in He_2 is much greater than in any other band spectrum. Moreover, since most of them correspond very closely with those of the He atom, it is possible to attach term designations to them with a certainty which is unattainable in any other case The helium band spectrum therefore offers a particularly favourable opportunity for the study of molecular electronic levels. Those identified up to the present are shown diagrammatically in fig. 6.

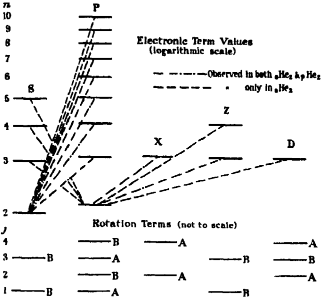


FIG. 6.—Electronic and Rotation Terms in He_2.

The term values and effective quantum numbers are given in Table II, the corresponding atomic quantum numbers being included for comparison.

Table II.—Electronic Terms and Effective Quantum Numbers in He_2

$_oHe_2$

	Total Quantum No.	2.	3.	4.	5	6
S	Electronic Term Value	34301·8	13893·1	7524·6	4713·2	
	Effective Quantum No.	1·788	2·810	3·818	4·824	
	Atomic Effective Qn. No.	*1·689*	*2·697*	*3·700*	*4·701*	
P	Electronic Term Value	29517·2	12794·5	7108·8	4516·3	3121·8
	Effective Quantum No.	1·928	2·928	3·928	4·928	5·927
	Atomic Effective Qn. No.	*1·937*	*2·933*	*3·932*	*4·931*	*5·931*
D	Electronic Term Value		12080·6			
	Effective Quantum No.		3·013			
	Atomic Effective Qn. No.		*2·997*			
X	Electronic Term Value		12533·0			
	Effective Quantum No.		2·958			
	Atomic Effective Qn. No.		—			
Z	Electronic Term Value		12734·0	7080		
	Effective Quantum No.		2·935	3·936		
	Atomic Effective Qn. No.		—	—		

Table II—(continued).

$_p\mathrm{He}_2$.

S	Electronic Term Value Effective Quantum No. Atomic Effective Qn. No.	31958·1 1·853 *1·850*				
P	Electronic Term Value Effective Quantum No Atomic Effective Qn. No.	28423·4 1·964 *2·009*	12481·4 2·964 *3·011*	6978·5 3·964 *4·011*	4448·4 4·965 *5·011*	3083·1 5·964 *6·011*
D	Electronic Term Value Effective Quantum No Atomic Effective Qn No		12040 3·018 *2·998*			
X	Electronic Term Value Effective Quantum No Atomic Effective Qn. No		12414·3 2·972 -			
Z	Electronic Term Value Effective Quantum No Atomic Effective Qn. No.		12580 2·953 --			

A striking feature of these results is that in every case the $_o\mathrm{He}_2$ value is less than the $_p\mathrm{He}_2$, and that a difference in the same sense exists between the ortho- and par-helium values. The inference would seem to be that the two atoms composing the molecule preserve their identity in combination, or in other words there is no "electron sharing" such as probably occurs in certain other molecules‡ (*e g*, BO, CN, CO) On this view He_2 would fall into line with other symmetrical molecules §

There is one important respect in which the atomic and molecular levels differ; there is nothing in the atomic scheme corresponding to the X and Z molecular levels. It is of great interest, therefore, to study their characteristics and to seek an explanation of them.

The most significant fact to be noted concerning them is that both are single, X having only A states and Z only B states This at once suggests that they may be associated in some way with the S level, the only other which is rotationally single (B states only) Another important consideration in determining level types is the minimum value of j which occurs. This is 0 for S levels, 1 for P, 2 for D, and so on. It is sometimes difficult in this case to draw conclusions from such evidence on account of the suppression of alternate rotation levels. Thus, for example, $j = 0$ is not observed in the S levels, but probably only because of the suppression of the even j levels. The P and D levels begin at $j = 1$ and $j = 2$ respectively in accordance with the above rule. In the case of the X level $j = 2$ is the lowest observed value, but it cannot be regarded as certain that $j = 0$ is absent, as the line P (1) to which it would give rise would be very weak and may have been missed on this account. The Z level is similar to the S in having $j = 1$ as the lowest observed value.

---

‡ See for example Mulliken, 'Phys. Rev.,' vol. 28, p. 494 (1926).

§ See Mecke, 'Z. Physik,' vol. 42, p 390 (1927).

It seems possible to account for most of the properties of the new levels on the lines of Hund's theory‡ of band spectra. In fact, the occurrence of molecular electronic levels additional to those of corresponding atoms is an important feature of the theory. In the case of the atom there is, neglecting multiplicity, a sequence of levels determined by each successive value (0, 1, 2, . . .) of l, the resultant electronic angular momentum. In the case of the diatomic molecule on this theory, the levels are not primarily determined by l, but by its component§ σ_l parallel to the nuclear axis. Thus in place of the three atomic S, P, D levels there ought to be six levels having $l = 0$, $\sigma_l = 0$ (S^S), $l = 1$, $\sigma_l = 0$ (S^P) or 1 (P^P); $l = 2$, $\sigma_l = 0$ (S^D), 1 (P^D) or 2 (D^D). The convenient notation used is that suggested by Mulliken.‖

If we identify, as seems natural from the missing lines and other evidence, the already known S, P, D levels with S^S, P^P, D^D above, the possibility of X and Z being S^P and S^D at once suggests itself. Both are rotationally single, as they would be on this supposition, and the states present are alternately odd and even, and A and B, in going from S to X to Z, thus :—S has 1_B, 3_B, 5_B, . ., X has 2_A, 4_A, 6_A, . . ., and Z has 1_B, 3_B, 5_B, ... It is not clear whether this can be definitely predicted from the theory at present, but it appears to harmonise with it in a general way. In any case the three are certainly related, and their natural order is certainly S, X, Z. Further evidence for this order is to be found in the progressive increase from S to Z in the complexity of the rotational term function. A similar tendency is also shown in going from S through P to D, but is less pronounced in that case. The absence of a Q branch in the X → P bands, and the peculiar intensity distribution in the Z → P, are the chief points which are not accounted for on this hypothesis. Since we have 2_A, 4_A, . X levels combining with 1_A, 2_B, 3_A, 4_B, . . P levels we should expect a Q_2 branch due to the transitions $2_A \rightarrow 2_B$, $4_A \rightarrow B_B$ . . .. Perhaps its absence may be due to the operation of a selection principle which excludes a Q branch if $\Delta l = 0$. This would not conflict with the presence of a Q branch in $Z \rightarrow P$ bands, since in that case $\Delta l = -1$ on the above hypothesis. It is not possible to derive any assistance from other band spectra, since the identifications of levels in them are much less reliable than here (see Hund's third paper referred to above).

---

‡ 'Z. Physik,' vol. 36, p. 657 (1926); vol. 40, p. 742 (1927); vol. 42, p. 93 (1927).

§ σ_l will be used here instead of Hund's i_l, in accordance with a suggestion of Prof Mulliken (private communication)

‖ 'Phys. Rev.,' vol. 29, p. 649 (1927).

On Hund's theory neither $\Delta l = 0$ nor $\Delta \sigma_l = 0$ is forbidden, so that, in general, many more electron transitions are to be expected in a molecule than in the corresponding atom. But in the present case the absence of alternate levels reduces the number very considerably. For example, X and S terms cannot combine, since the j values are respectively 2_A, 4_A, . . . and 1_B, 3_B, . . . Q branches are obviously impossible and P, R are precluded by the "crossing over" ($A \longleftrightarrow B$) which must necessarily occur. Detailed consideration of this nature leads to the conclusion that only the following should be possible :—

$$S^S \longleftrightarrow P^P,\ S^P \longleftrightarrow P^P,\ S^D \longleftrightarrow P^P,\ P^P \longleftrightarrow P^D,\ P^P \longleftrightarrow D^D.$$

If we identify S^P with X and S^D with Z as proposed, all of these have been observed except $P^P \longleftrightarrow P^D$, the latter level not having been identified as yet. There is, however, a well-marked band of peculiar structure in the neighbourhood of λ 5800 which may possibly turn out to be one of this series. Another characteristic of the new levels which should be of assistance in discovering their nature is the lowest value of the electronic (total) quantum number which occurs. For an S^P level this should be 2 and for an S^D 3. It is therefore of importance to determine whether the 2X and 2Z levels exist. If so, they would give rise to bands $3P \rightarrow 2X$ and $3P \rightarrow 2Z$ near λ 6320 and λ 6050 respectively, both estimates being liable to considerable error on account of the extrapolation involved. In the former case there are some lines unaccounted for in the λ 6240 region which might belong to the band in question. The λ 6050 region is very complex, and has not yet been analysed. In neither case would it be permissible to conclude that the predicted band is absent.

It is probable that further evidence of value would emerge from a more detailed and comprehensive study of the Zeeman effect in this spectrum. It would be of particular interest to see what kind of effect is shown by the $D \rightarrow P$ bands, and whether this bears any relation to that observed in the case of the $4Z \rightarrow 2P$.[‡] Unfortunately the equipment at present at our disposal does not permit of our undertaking this.

Summary.

The details and analysis of five new He_2 bands are given. One of these is a weak vibrational band associated with the known band near λ 5730 ($3D \rightarrow 2P$

[‡] Curtis and Jevons, *loc. cit.*

of $_o\mathrm{He}_2$). Another is the $_p\mathrm{He}_2$ counterpart of the $_o\mathrm{He}_2$ band designated 3X → 2P and described in the preceding paper of this series. The remaining three have 2P as the final electronic level and a new type of level (Z) as initial. It is rotationally single, like S and X, but the rotation terms cannot be represented by the usual type of formula, nor are the relative intensities of the branches at all similar to those in other bands. In consequence of the abnormal character of the initial level the appearance of the Z → P bands is very peculiar; the wave-numbers of the R branch, for example, decrease continuously with increase of rotational quantum number, thus giving it the appearance of a P branch. Two perturbations are recorded in these bands, one a large displacement and the other a splitting into two components of about equal intensity. The Zeeman effect has already been found to be very unusual in magnitude and character for the band 4Z → 2P.

The X and Z levels are clearly additional to the ordinary atomic system of levels, and evidence is discussed which leads to a tentative identification of them with certain new types predicted by Hund for diatomic molecules, but not hitherto definitely established by observation. Although the chief properties of the new bands may readily be accounted for on this view, several unexplained peculiarities remain, such as the absence of Q branches in X → P transitions and the relative intensities of the branches in Z → P transitions.

[*Note added, August* 6.—After the above was completed we received the current number of the 'Zeitschrift f. Physik' (vol. 49, pp. 637–669) containing a paper by Takamine, Dieke and Suga, which deals with the bands here described, as well as with those of IV and some others. It is on rather different lines, relating primarily to the presentation of the data, which are mainly in satisfactory agreement with ours. The chief differences are: (1) they find a weak and somewhat doubtful Q branch in 3X → 2P of $_o\mathrm{He}_2$, and traces of a possible Q branch in 3X → 2P of $_p\mathrm{He}_2$. If these are genuine they are probably too weak to show on our plates. (2) They find only two branches in all the Z → 2P bands, and these (our Q and R) they interpret as P and Q. Our identification of a third (P) branch in one case affords strong evidence in favour of our view, but further work is desirable in connection with both these questions.]

Fluid Motion in a Curved Channel.

By W. R. Dean, M.A., Imperial College of Science.

(Communicated by S. Chapman, F.R.S.—Received July 31, 1928.)

Experimental work due to Prof. J. Eustice* has shown that there is no marked critical velocity for a fluid flowing through a curved pipe. If the pipe is straight there is a sudden increase in the loss of head as soon as the velocity exceeds its critical value; below the critical the loss of head varies as the first power of the velocity, but above it approximately as the second power. But if the pipe is curved there does not appear to be such a sudden change at any velocity of flow. One possibility is that flow through a curved pipe is stream-line at velocities much greater than the critical for a straight pipe, but experiment† seems to show that the critical velocity is smaller in a curved pipe than in a straight one. If then the motion in a curved pipe becomes unstable at a velocity somewhat less than the critical for a straight pipe, the absence of a sudden increase in the loss of head in this region suggests that the stream-line motion in a curved pipe (unlike that in a straight pipe) is unstable for small disturbances. A similar problem shows that it is not unlikely that curvature may have such an effect: it is believed that uniform shearing motion between flat plates is stable for small disturbances,‡ but Prof. G. I. Taylor§ has shown that shearing motion between concentric cylinders can in certain conditions become unstable for small disturbances.

A theoretical investigation of the stability of flow in a curved pipe is certain to be a matter of great difficulty, and therefore a simplified form of the problem, the stability of flow under pressure through a curved channel (*i.e.*, between concentric cylinders), is here considered. It is shown that the motion can become unstable for a small disturbance of exactly the type found by Taylor to be possible in motion between rotating cylinders.

It is assumed that d, the distance between the two cylinders, is small in comparison with a, the smaller radius, and if this assumption is introduced at the start the equations, which are complicated when exact, become quite simple. As a check on the method seemed desirable, the results it gives are shown to be approximately the same as those found by Taylor by a more satisfactory procedure in which an equivalent assumption is made only after an exact solution of the equations has been obtained.

---

* 'Roy. Soc. Proc.,' A, vol. 84, p. 107 (1910).

† J. Eustice, 'Roy. Soc. Proc.,' A, vol. 85, p. 119 (1911).

‡ *E.g.*, R. V. Southwell, 'Phil. Mag.,' vol. 48, p. 545 (1924).

§ 'Phil. Trans.,' A, vol. 223, p. 289.

The critical velocity for the type of disturbance considered is found to be given by

$$\bar{v}d/\nu = 36\,(a/d)^{\frac{1}{2}},$$

where $\bar{v}$ is the mean velocity.

This value is small enough to suggest that the disturbance may (in a certain range of values of a/d) be that which actually arises when stream-line motion breaks down. This, however, could hardly be decided except by experiment, since it would be almost impossible to examine theoretically all possible types of small disturbance. What the work of the paper does show is that a type of small disturbance which could not persist in a straight channel is possible in a curved channel, and this makes more likely the explanation suggested above of the absence of sudden changes when the velocity of fluid flowing through a curved pipe passes through its critical value.

2. The equations for the steady motion of incompressible fluid, referred to cylindrical co-ordinates (r, θ, z), are, if the velocity components associated with these co-ordinates are independent of θ,

$$\mathrm{U}\frac{\partial \mathrm{U}}{\partial r} + \mathrm{W}\frac{\partial \mathrm{U}}{\partial z} - \frac{\mathrm{V}^2}{r} = -\frac{\partial}{\partial r}\left(\frac{\mathrm{P}}{\rho}\right) + \nu\left(\frac{\partial^2 \mathrm{U}}{\partial r^2} + \frac{1}{r}\frac{\partial \mathrm{U}}{\partial r} + \frac{\partial^2 \mathrm{U}}{\partial z^2} - \frac{\mathrm{U}}{r^2}\right), \quad (1)$$

$$\mathrm{U}\frac{\partial \mathrm{V}}{\partial r} + \mathrm{W}\frac{\partial \mathrm{V}}{\partial z} + \frac{\mathrm{UV}}{r} = -\frac{1}{r}\frac{\partial}{\partial \theta}\left(\frac{\mathrm{P}}{\rho}\right) + \nu\left(\frac{\partial^2 \mathrm{V}}{\partial r^2} + \frac{1}{r}\frac{\partial \mathrm{V}}{\partial r} + \frac{\partial^2 \mathrm{V}}{\partial z^2} - \frac{\mathrm{V}}{r^2}\right), \quad (2)$$

and

$$\mathrm{U}\frac{\partial \mathrm{W}}{\partial r} + \mathrm{W}\frac{\partial \mathrm{W}}{\partial z} = -\frac{\partial}{\partial z}\left(\frac{\mathrm{P}}{\rho}\right) + \nu\left(\frac{\partial^2 \mathrm{W}}{\partial r^2} + \frac{1}{r}\frac{\partial \mathrm{W}}{\partial r} + \frac{\partial^2 \mathrm{W}}{\partial z^2}\right), \quad (3)$$

where P is the pressure, ρ the density and ν the kinematic viscosity. The equation of continuity is

$$\frac{\partial \mathrm{U}}{\partial r} + \frac{\mathrm{U}}{r} + \frac{\partial \mathrm{W}}{\partial z} = 0. \quad (4)$$

Suppose that the fluid is contained between the cylinders $r = a$ and $r = a + d$, and that

$$\mathrm{P}/\rho = k\theta + f(r),$$

where k is constant; the pressure then varies across a section of the channel but falls steadily (if $k<0$) along the central surface of the channel, $r = a + d/2$, as θ increases.

There is an exact solution of the equations appropriate to this case, which we can write

$$U = W = 0, \quad V = V_0, \quad p = P_0;$$

the actual expressions for V_0 and P_0 are not required.

Let $r = a + x$, so that the cylinders are $x = 0$ and $x = d$; then a first approximation to V_0 is given by

$$V_0 = \frac{k}{2\nu a}(x^2 - xd), \tag{5}$$

if terms of the relative order d/a are neglected.

3 Consider now a steady motion slightly different from that above:

$$U = u, \quad V = V_0 + v, \quad W = w, \quad P = P_0 + p;$$

u, v, w and p are all assumed to be small and independent of θ.

Substituting these expressions in equations (1) to (4), using the relation between P_0 and V_0 and ignoring squares and products of u, v, w, we reach the equations

$$-\frac{2V_0 v}{a+x} = -\frac{\partial}{\partial x}\left(\frac{p}{\rho}\right) + \nu\left\{\frac{\partial^2 u}{\partial x^2} + \frac{1}{a+x}\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial z^2} - \frac{u}{(a+x)^2}\right\}, \tag{6}$$

$$u\left(\frac{\partial V_0}{\partial x} + \frac{V_0}{a+x}\right) = \nu\left\{\frac{\partial^2 v}{\partial x^2} + \frac{1}{a+x}\frac{\partial v}{\partial x} + \frac{\partial^2 v}{\partial z^2} - \frac{v}{(a+x)^2}\right\}, \tag{7}$$

$$0 = -\frac{\partial}{\partial z}\left(\frac{p}{\rho}\right) + \nu\left\{\frac{\partial^2 w}{\partial x^2} + \frac{1}{a+x}\frac{\partial w}{\partial x} + \frac{\partial^2 w}{\partial z^2}\right\}, \tag{8}$$

and

$$0 = \frac{\partial u}{\partial x} + \frac{u}{a+x} + \frac{\partial w}{\partial z}. \tag{9}$$

4. Now let the assumption that d/a is small be introduced. A first approximation to the equations is then obtained by neglecting such terms as

$$\frac{1}{a+x}\frac{\partial u}{\partial x} - \frac{u}{(a+x)^2},$$

in comparison with $\frac{\partial^2 u}{\partial x^2}$. Again $-2V_0 v/(a+x)$ can be written $-2V_0 v/a$, but we have evidently no grounds for neglecting the latter term in comparison with $\nu\frac{\partial^2 u}{\partial x^2}$.

With these approximations, equations (6) to (9) become

$$-\frac{2V_0 v}{a} = -\frac{\partial}{\partial x}\left(\frac{p}{\rho}\right) + \nu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}\right), \tag{10}$$

$$u\frac{\partial V_0}{\partial x} = \nu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2}\right), \tag{11}$$

$$0 = -\frac{\partial}{\partial z}\left(\frac{p}{\rho}\right) + \nu\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial z^2}\right), \tag{12}$$

$$0 = \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}. \tag{13}$$

We have to find the condition that there may be a non-zero solution of these equations subject to the boundary conditions

$$u = v = w = 0\,; \quad x = 0,\, d \tag{14}$$

5 Although it is fairly clear that the most important terms in equations (6) to (9) are retained in equations (10) to (13), the process of approximation is not on a precise basis. As it seemed desirable to have a check on the method, it is now in §§ 5–7 applied to the motion of fluid between rotating cylinders, since this problem has been worked out by Prof. G. I. Taylor starting from an exact solution of equations very similar to those above.

Equations (6) to (9) apply to this new problem provided we write

$$V_0 = Ar + B/r = A(a + x) + B/(a + x),$$

where A and B are constants whose values depend on the angular velocities of the two cylinders.

In this case, unlike the last, V_0 does not vanish at the boundaries, and we have therefore no reason for neglecting $V_0/(a + x)$ in comparison with $\partial V_0/\partial x$; in fact, since the expression for V_0 is simple there is no point in making any approximation in the first instance in the left-hand sides of (6) and (7). Making the other approximations as before, we get

$$-2\{A + B/(a + x)^2\}\,v = -\frac{\partial}{\partial x}\left(\frac{p}{\rho}\right) + \nu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}\right),$$

$$2Au = \nu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2}\right),$$

$$0 = -\frac{\partial}{\partial z}\left(\frac{p}{\rho}\right) + \nu\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial z^2}\right),$$

$$0 = \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}.$$

Assume now that u and v are proportional to $\cos \lambda z$ and w to $\sin \lambda z$; and eliminate p from the equations. There results:

$$\nu \left(\frac{\partial^2}{\partial x^2} - \lambda^2\right) v = 2\mathrm{A}u, \tag{15}$$

$$\frac{\nu}{\lambda}\left(\frac{\partial^2}{\partial x^2} - \lambda^2\right)\left(\frac{\partial w}{\partial x} + \lambda u\right) = -2\{\mathrm{A} + \mathrm{B}/(a+x)^2\} v, \tag{16}$$

and

$$\frac{\partial u}{\partial x} + \lambda w = 0. \tag{17}$$

u, v and w are now functions of x only; boundary conditions (14) are unchanged.

6. Let

$$u = b_1 \sin \frac{\pi x}{d} + b_2 \sin \frac{2\pi x}{d} + \dots;$$

this form will satisfy the condition $u = 0$ when $x = 0, d$. From (15)

$$v = -\frac{2\mathrm{A}}{\nu} \Sigma \frac{b_m}{\lambda^2 + m^2\pi^2/d^2} \sin \frac{m\pi x}{d},$$

no arbitrary constants being needed since v also must vanish when $x = 0, d$.

Then

$$-2\{\mathrm{A} + \mathrm{B}/(a+x)^2\} v$$
$$= \frac{4\mathrm{A}}{\nu}\left\{\mathrm{A} + \frac{\mathrm{B}}{a^2}\left(1 - \frac{2x}{a} + \frac{3x^2}{a^2}\right)\right\} \Sigma \frac{b_m}{\lambda^2 + m^2\pi^2/d^2} \sin \frac{m\pi x}{d};$$

an approximation is here made in the coefficient of v.

Now let

$$\left\{\mathrm{A} + \frac{\mathrm{B}}{a^2}\left(1 - \frac{2x}{a} + \frac{3x^2}{a^2}\right)\right\} \sin \frac{m\pi x}{d} = \sum_{r=1}^{\infty} a_{mr} \sin \frac{r\pi x}{d};$$

then

$$a_{mm} = \left(\mathrm{A} + \frac{\mathrm{B}}{a^2}\right) - \frac{\mathrm{B}d}{a^3} + \frac{\mathrm{B}d^2}{a^4}\left(1 - \frac{3}{2m^2\pi^2}\right),$$

$$a_{mn} = \frac{\mathrm{B}d^2}{\pi^2 a^4} \frac{24mn}{(m^2 - n^2)^2}, \quad (m+n) \text{ even},$$

$$= \frac{\mathrm{B}d}{\pi^2 a^3} \frac{16mn}{(m^2 - n^2)^2}\left(1 - \frac{3d}{2a}\right), \quad (m+n) \text{ odd}.$$

We now have

$$-2\left\{\mathrm{A} + \frac{\mathrm{B}}{a^2}\left(1 - \frac{2x}{a} + \frac{3x^2}{a^2}\right)\right\} v$$

$$= \frac{4\mathrm{A}}{\nu} \sum_{m=1}^{\infty} \frac{b_m}{\lambda^2 + m^2\pi^2/d^2}\left\{a_{m1} \sin \frac{\pi x}{d} + a_{m2} \sin \frac{2\pi x}{d} + \dots\right\}$$

$$= \sum_{=1}^{\infty} c_r \sin \frac{r\pi x}{d}, \tag{18}$$

and we can write equation (16)

$$\frac{\nu}{\lambda}\left(\frac{\partial^2}{\partial x^2}-\lambda^2\right)\left(\frac{\partial w}{\partial x}+\lambda u\right)=\Sigma\, c_r \sin\frac{r\pi x}{d}.$$

This gives

$$\frac{\nu}{\lambda}\left(\frac{\partial w}{\partial x}+\lambda u\right)=C'e^{\lambda x}+D'e^{-\lambda x}-\Sigma\frac{c_r}{\lambda^2+r^2\pi^2/d^2}\sin\frac{r\pi x}{d},$$

$$\frac{\nu}{\lambda}\frac{\partial w}{\partial x}=C'e^{\lambda x}+D'e^{-\lambda x}-\Sigma\sin\frac{r\pi x}{d}\left\{\frac{c_r}{\lambda^2+r^2\pi^2/d^2}+\nu b_r\right\},$$

and finally

$$\nu w/\lambda=E+Ce^{\lambda x}+De^{-\lambda x}+\sum_{r=1}^{\infty}\frac{d}{r\pi}\cos\frac{r\pi x}{d}\left\{\frac{c_r}{\lambda^2+r^2\pi^2/d^2}+\nu b_r\right\}.$$

Next expand the complementary function as a cosine series. Before doing so it is convenient to change the arbitrary constants; let

$$E+Ce^{\lambda x}+De^{-\lambda x}=E_1+\frac{dC_1\cosh\{\lambda(x-d/2)\}}{4\lambda\sinh(\lambda d/2)}-\frac{dD_1\sinh\{\lambda(x-d/2)\}}{4\lambda\cosh(\lambda d/2)}, \tag{19}$$

where E_1, C_1 and D_1 are new arbitrary constants. If the right-hand side is written as the Fourier Series

$$e_0+e_1\cos\frac{\pi x}{d}+e_2\cos\frac{2\pi x}{d}+\ldots,$$

it can be seen that

$$e_m=\frac{C_1}{\lambda^2+m^2\pi^2/d^2},\ m\text{ even},\quad -\frac{D_1}{\lambda^2+m^2\pi^2/d^2},\ m\text{ odd}. \tag{20}$$

We can now write

$$\nu w/\lambda=e_0+\sum_{r=1}^{\infty}\cos\frac{r\pi x}{d}\left[\frac{d}{r\pi}\left\{\frac{c_r}{\lambda^2+r^2\pi^2/d^2}+\nu b_r\right\}+e_r\right].$$

There remain to be satisfied equation (17) and the boundary condition $w=0$, $x=0$, d.

The latter conditions become, using (17),

$$\left.\begin{aligned}b_1+2b_2+3b_3+\ldots&=0\\-b_1+2b_2-3b_3+\ldots&=0\end{aligned}\right\}, \tag{21}$$

while equation (17) is satisfied provided $e_0=0$ and

$$\frac{d}{r\pi}\left\{\frac{c_r}{\lambda^2+r^2\pi^2/d^2}+\nu b_r\right\}+e_r+\frac{r\pi\nu}{d\lambda^2}b_r=0,\quad r=1,2,\ldots. \tag{22}$$

The condition that there may be a set of non-zero values of C_1, D_1, b_1, b_2, ... satisfying these equations can now be written down by equating to zero an infinite determinant.

7. It remains to show that the equation so derived is approximately the same as that given by Taylor. The latter* is

$$\begin{vmatrix} 0 & 0 & (1+d^2\lambda^2/\pi^2) & 2\,(4+d^2\lambda^2/\pi^2) & 3\,(9+d^2\lambda^2/\pi^2) & \cdots \\ 0 & 0 & -(1+d^2\lambda^2/\pi^2) & 2\,(4+d^2\lambda^2/\pi^2) & -3\,(9+d^2\lambda^2/\pi^2) & \cdots \\ 1 & -1 & \mathrm{L}_1' & {}_1c_2 & {}_1c_3 & \cdots \\ 2 & 2 & {}_2c_1 & \mathrm{L}_2' & {}_2c_3 & \cdots \\ 3 & -3 & {}_3c_1 & {}_3c_2 & \mathrm{L}_3' & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix} = 0, \quad (23)$$

where

$$ {}_nc_m = \frac{8mn\,\Omega_1\,(1-\mu)}{(m^2-n^2)^2\,\pi^2}, \quad (m+n) \text{ odd,}$$

$$ = \frac{8mn\,\Omega_1\,(1-\mu)}{(m^2-n^2)^2\,\pi^2}\left(\frac{3d}{2a}\right), \ (m+n) \text{ even,}$$

$$\mathrm{L}_m' = \frac{\nu^2}{4\mathrm{A}\lambda^2}(\lambda^2+m^2\pi^2/d^2)^3 + \frac{\Omega_1}{2}(1+\mu) - \Omega_1\,(1-\mu)\frac{d}{a}\left(\frac{1}{4}+\frac{3}{4m^2\pi^2}\right)$$

Ω_1, Ω_2 are the angular velocities of the inner and outer cylinders, respectively, and $\mu = \Omega_2/\Omega_1$.

The determinant obtained from equations (21) and (22) requires some manipulation before it can be set in the form of that in (23). Equation (22) can be written

$$c_r + \frac{r\pi}{d}(\lambda^2 + r^2\pi^2/d^2)\,c_r + \frac{\nu b_r}{\lambda^2}(\lambda^2 + r^2\pi^2/d^2)^2 = 0,$$

while from (18)

$$c_r = \frac{4\mathrm{A}}{\nu}\sum_{m=1}^{\infty}\frac{b_m a_{mr}}{\lambda^2 + m^2\pi^2/d^2}.$$

Hence the first two equations of the set ($r = 1, 2$) are

$$\frac{\pi}{d}\mathrm{D}_1 + b_1\left\{\frac{\nu}{\lambda^2}(\lambda^2+\pi^2/d^2)^2 + \frac{4\mathrm{A}}{\nu}\frac{a_{11}}{\lambda^2+\pi^2/d^2}\right\} + b_2\frac{4\mathrm{A}}{\nu}\frac{a_{21}}{\lambda^2+4\pi^2/d^2} + \ldots = 0,$$

and

$$\frac{2\pi}{d}\mathrm{C}_1 + b_1\frac{4\mathrm{A}}{\nu}\frac{a_{12}}{\lambda^2+\pi^2/d^2} + b_2\left\{\frac{\nu}{\lambda^2}(\lambda^2+4\pi^2/d^2)^2 + \frac{4\mathrm{A}}{\nu}\frac{a_{22}}{\lambda^2+4\pi^2/d^2}\right\} + \ldots = 0.$$

---

* *Loc. cit.*, p. 308, equation (5.40); the notation has been altered in some places so as to be nearer that of this paper.

The condition can now be written

$$\left|\begin{array}{cccccc} 0 & 0 & 1 & 2 & . \\ 0 & 0 & -1 & 2 & \cdots \\ 1 & 0 & \dfrac{\nu}{\lambda^2}(\lambda^2+\pi^2/d^2)^2+\dfrac{4\mathrm{A}}{\nu}\dfrac{a_{11}}{\lambda^2+\pi^2/d^2} & \dfrac{4\mathrm{A}}{\nu}\dfrac{a_{21}}{\lambda^2+4\pi^2/d^2} & . \\ 0 & 2 & \dfrac{4\mathrm{A}}{\nu}\dfrac{a_{12}}{\lambda^2+\pi^2/d^2} & \dfrac{\nu}{\lambda^2}(\lambda^2+4\pi^2/d^2)^2+\dfrac{4\mathrm{A}}{\nu}\dfrac{a_{22}}{\lambda^2+4\pi^2/d^2}, \\ 3 & 0 & \dfrac{4\mathrm{A}}{\nu}\dfrac{a_{13}}{\lambda^2+\pi^2/d^2} & \dfrac{4\mathrm{A}}{\nu}\dfrac{a_{23}}{\lambda^2+4\pi^2/d^2} & \cdots \\ . & . & . & . & . \end{array}\right| = 0.$$

Some simple reductions bring this equation to the form (23).

Multiply all rows except the first two by $\nu/4\mathrm{A}$, and then divide by this factor the first two columns; then multiply the third column by $(\lambda^2+\pi^2/d^2)$, the fourth by $(\lambda^2+4\pi^2/d^2)$, and so on; lastly multiply the first two rows by d^2/π^2.

Then the condition is

$$\left|\begin{array}{ccccc} 0 & 0 & 1+d^2\lambda^2/\pi^2 & 2\,(4+d^2\lambda^2/\pi^2) & \cdots \\ 0 & 0 & -(1+d^2\lambda^2/\pi^2) & 2\,(4+d^2\lambda^2/\pi^2) & \cdots \\ 1 & 0 & \dfrac{\nu^2}{4\mathrm{A}\lambda^2}(\lambda^2+\pi^2/d^2)^3+a_{11} & a_{21} & \cdots \\ 0 & 2 & a_{12} & \dfrac{\nu^2}{4\mathrm{A}\lambda^2}(\lambda^2+4\pi^2/d^2)^3+a_{22}, & \cdots \\ & . & . & . & \cdots \end{array}\right| = 0. \qquad (24)$$

The first two columns can evidently be reduced to the first two columns in (23), and it therefore remains to show that

$$a_{m} = {}_{n}c_{mn},$$

$$a_{mm} = \frac{\Omega_1}{2}(1+\mu) - \Omega_1(1-\mu)\frac{d}{a}\left(\tfrac{1}{4}+\frac{3}{4m^2\pi^2}\right).$$

If $(m+n)$ is odd, $a_{mn} = {}_{n}c_{m}$ provided

$$\Omega_1(1-\mu) = \frac{2\mathrm{B}d}{a^3}\left(1-\frac{3d}{2a}\right).$$

Now

$$\Omega_1(1-\mu) = \Omega_1 - \Omega_2 = \{A + B/a^2\} - \{A + B/(a+d)^2\}$$
$$= \frac{2Bd}{a^3}\left(1 - \frac{3d}{2a} + \frac{2d^2}{a^2} - \ldots\right).$$

Hence if terms of the relative order d^2/a^2 are ignored the equality holds.

If $(m+n)$ is even, $a_{mn} = {}_nc_m$ provided

$$\Omega_1(1-\mu)\frac{3d}{2a} = \frac{3Bd^2}{a^4},$$

and this is true to the same accuracy as before.

Lastly,

$$\frac{\Omega_1}{2}(1+\mu) - \Omega_1(1-\mu)\frac{d}{a}\left(\frac{1}{4} + \frac{3}{4m^2\pi^2}\right) = \left(A + \frac{B}{a^2}\right) - \frac{Bd}{a^3} + \frac{Bd^2}{a^4}\left(1 - \frac{3}{2m^2\pi^2}\right),$$

if terms of higher order in d/a are ignored, and the expression on the right is equal to a_{mm}.

Equations (23) and (24) are therefore approximately the same, and we can conclude that in this case the process of approximation is justifiable. This will be assumed for the problem treated in the present paper.

8. If p is eliminated from equations (10) to (13), and it is assumed that u and v are proportional to $\cos \lambda z$ and w to $\sin \lambda z$, the equations become

$$\left(\frac{\partial^2}{\partial x^2} - \lambda^2\right)\left(\frac{\partial w}{\partial x} + \lambda u\right) = -\frac{2V_0\lambda}{\nu a}v,$$

$$\left(\frac{\partial^2}{\partial x^2} - \lambda^2\right)v = \frac{1}{\nu}\frac{\partial V_0}{\partial x}u,$$

and

$$\frac{\partial u}{\partial x} + \lambda w = 0. \tag{25}$$

Substituting in the first two equations the approximate expression for V_0 given by (5), we have

$$\left(\frac{\partial^2}{\partial x^2} - \lambda^2\right)\left(\frac{\partial w}{\partial x} + \lambda u\right) = -\frac{k\lambda}{\nu^2 a^2}(x^2 - xd)v, \tag{26}$$

and

$$\left(\frac{\partial^2}{\partial x^2} - \lambda^2\right)v = \frac{k}{2\nu^2 a}(2x - d)u. \tag{27}$$

The ratio of the coefficient of v in (26) to the coefficient of u in (27) is of order d/a, and it may appear at first sight that the former term should therefore be neglected. It will, however, be seen that these two coefficients are in effect multiplied together, so that to obtain a first approximation both terms must be kept in; further, although we are assuming that d/a is small, it does not necessarily follow that (roughly speaking) $V_0 d/a$ is small.

We have now to find the condition for a non-zero solution of equations (25) to (27) subject to the boundary condition (14).

9. It is convenient to write the equations in non-dimensional form.

Let

$$y = \pi x/d, \quad l = d\lambda/\pi ,$$

then the equations become

$$\left(\frac{\partial^2}{\partial y^2} - l^2\right)\left(\frac{\partial w}{\partial y} + lu\right) = B\left(\frac{y^2}{\pi^2} - \frac{y}{\pi}\right) lv, \tag{28}$$

$$\left(\frac{\partial^2}{\partial y^2} - l^2\right) v = - A\left(\frac{2y}{\pi} - 1\right) u, \tag{29}$$

and

$$\frac{\partial u}{\partial y} + lw = 0, \tag{30}$$

where

$$A = - kd^3/2\pi^2\nu^2 a, \quad B = - kd^4/\pi^2\nu^2 a^2. \tag{31}$$

The boundary conditions are

$$u = v = w = 0 ; \quad y = 0, \pi. \tag{32}$$

In the solution of these equations two such expansions as that shown in (18) have successively to be made; further, we have nominally to find the values of A and B for all values of l, so that the least possible values of A and B leading to a non-zero solution of the equations can be found. As this would involve rather heavy numerical work a preliminary exploration seemed desirable. For this purpose $(2y/\pi - 1)$ in equation (29) was replaced by $- \cos y$, and $(y^2/\pi^2 - y/\pi)$ in (28) by $- (1 - \cos 2y)/8$; the two expansions required were then, of course, very simple. It was found that AB had a minimum value of roughly 860 when l^2 was about 1·5. This result served as a guide in the arithmetical work of the exact solution.

10. In solving equations (28) to (30) we proceed exactly as in §§ 5–7.

Let

$$u = a_1 \sin y + a_2 \sin 2y + \dots , \tag{33}$$

$$(2y/\pi - 1)\, u = b_1 \sin y + b_2 \sin 2y + \dots ; \tag{34}$$

then from (29)

$$v = A \Sigma \frac{b_m}{l^2 + m^2} \sin my, \tag{35}$$

no arbitrary constants being required since v must vanish when $y = 0$ and when $y = \pi$.

Further, let

$$\mathrm{B}\,(y^3/\pi^2 - y/\pi)\, lw = \mathrm{AB}l\,\Sigma\, c_m \sin my\,;$$

then from (28)

$$\frac{\partial w}{\partial y} + lu = \alpha' e^{ly} + \beta' e^{-ly} - \mathrm{AB}l\,\Sigma \frac{c_m}{l^2 + m^2} \sin my,$$

so that

$$\frac{\partial w}{\partial y} = \alpha' e^{ly} + \beta' e^{-ly} - \Sigma \sin my \left(la_m + \frac{\mathrm{AB}l}{l^2 + m^2} c_m\right)$$

and

$$w = \alpha e^{ly} + \beta e^{-ly} + \gamma + \Sigma \frac{\cos my}{m} \left(la_m + \frac{\mathrm{AB}l}{l^2 + m^2} c_m\right).$$

The complementary function must be expanded as a cosine series, and as before in § 6 it is convenient to alter the arbitrary constants; equations (19) and (20) can be used here if l is substituted for λ and π for d.

Then

$$w = e_0 + \Sigma \cos my \left[\frac{1}{m}\left\{la_m + \frac{\mathrm{AB}l}{l^2 + m^2} c_m\right\} + e_m\right],$$

where

$$e_m = \frac{\mathrm{C}}{l^2 + m^2},\ m \text{ even}, \qquad = \frac{\mathrm{D}}{l^2 + m^2},\ m \text{ odd},$$

C, D and e_0 being arbitrary constants.

Equation (30) and the boundary condition $w = 0$, $y = 0$, π, are satisfied if

$$a_1 + 2a_2 + 3a_3 + \ldots = 0,$$

$$-a_1 + 2a_2 - 3a_3 + \ldots = 0,$$

and

$$\frac{1}{m}\left\{la_m + \frac{\mathrm{AB}l}{l^2 + m^2} c_m\right\} + e_m + \frac{ma_m}{l} = 0, \quad m = (1, 2, \ldots).$$

The first two equations of the set above can be written

$$a_1 (l^2 + 1)^2 + \mathrm{AB}l^2 c_1 + l\mathrm{D} = 0,$$

$$a_2 (l^2 + 4)^2/2 + \mathrm{AB}l^2 c_2/2 + l\mathrm{C} = 0$$

If then

$$c_m = \sum_{r=1}^{\infty} \alpha_{mr} a_r,$$

the condition for a significant solution of the equations is

$$\begin{vmatrix} 0 & 0 & 1 & 0 & 3 & \cdots \\ 0 & 0 & 0 & 2 & 0 & \cdots \\ -1 & 0 & C_1 + X\alpha_{11} & X\alpha_{12} & X\alpha_{13} & \cdots \\ 0 & -1 & \dfrac{X\alpha_{21}}{2} & C_2 + \dfrac{X\alpha_{22}}{2} & \dfrac{X\alpha_{23}}{2} & . \\ -1 & 0 & \dfrac{X\alpha_{31}}{3} & \dfrac{X\alpha_{32}}{3} & C_3 + \dfrac{X\alpha_{33}}{3} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix} = 0, \quad (36)$$

where

$$X = ABl^2, \quad C_m = (l^2 + m^2)^2/m.$$

Now divide the third column of the determinant by C_1, the fourth by C_2, and so on; then add to the first column the sum of the third, fifth, .. columns, and add to the second column the sum of the fourth, sixth, ... columns. There results

$$\begin{vmatrix} \Sigma\left(\dfrac{1}{C_1}\right) & 0 & \dfrac{1}{C_1} & 0 & \cdots \\ 0 & \Sigma\left(\dfrac{2}{C_2}\right) & 0 & \dfrac{2}{C_2} & \cdots \\ \dfrac{X\alpha_{11}}{C_1} + \dfrac{X\alpha_{13}}{C_3} + \dots, & \dfrac{X\alpha_{12}}{C_2} + \dfrac{X\alpha_{14}}{C_4} + & 1 + \dfrac{X\alpha_{11}}{C_1} & \dfrac{X\alpha_{12}}{C_2} & . \\ \dfrac{X\alpha_{21}}{2C_1} + \dfrac{X\alpha_{23}}{2C_3} + \dots, & \dfrac{X\alpha_{22}}{2C_2} + \dfrac{X\alpha_{24}}{2C_4} + & \dfrac{X\alpha_{21}}{2C_1} & 1 + \dfrac{X\alpha_{22}}{2C_2} & . \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix} = 0. \quad (37)$$

11. We have now to find expressions for the coefficients α.

$$\frac{\pi}{2} c_m = \int_0^\pi \left(\frac{y^2}{\pi^2} - \frac{y}{\pi}\right)\left(\sum_{(n)} \frac{b_n}{l^2 + n^2} \sin ny\right) \sin my \, dy,$$

and it can be seen that

$$c_m = \frac{2}{\pi^2}\Bigg[-\frac{b_m}{l^2 + m^2}\left(\frac{\pi^2}{12} + \frac{1}{4m^2}\right) + \Sigma' \frac{b_n}{l^2 + n^2}\left\{\frac{1 + \cos(m-n)\pi}{2(m-n)^2} - \frac{1 + \cos(m+n)\pi}{2(m+n)^2}\right\}\Bigg]; \quad (38)$$

the summation sign is accented to show that the term given by $n = m$ is to be omitted.

Again,

$$\frac{\pi}{2} b_m = \int_0^\pi \left(\frac{2y}{\pi} - 1\right)\left(\sum_{(n)} a_n \sin ny\right) \sin my \, dy,$$

so that

$$b_m = -\frac{4}{\pi^2} \Sigma' a_n \left\{ \frac{1 - \cos(m-n)\pi}{2(m-n)^2} - \frac{1 - \cos(m+n)\pi}{2(m+n)^2} \right\}, \qquad (39)$$

where the accent has the same significance as before.

The numerical coefficients in equations (38) and (39) must be calculated; then the series for any α can be written down at once. It can be shown for instance that

$$\alpha_{12} = \frac{8}{\pi^4}\left[\frac{0 \cdot 9533}{l^2+1} - \frac{0 \cdot 1800}{l^2+9} - \frac{0 \cdot 0031}{l^2+25} - \frac{0 \cdot 0003}{l^2+49} - \dots\right].$$

For a given value of l any term in the determinant can be calculated. If $(m+n)$ is even $\alpha_{mn} = 0$. For α_{11}, for instance, is the coefficient of a_1 in c_1, and, therefore, the coefficient of a_1 in a series containing $b_1, b_3, b_5 \dots$, ; from equation (39) it appears that the coefficient is zero.

The series for the coefficients α converge rapidly. The series for α_{16} converges a little more slowly than that for α_{12}, but, on the other hand, the former has subsequently to be divided by C_6 and the latter by C_2. In no case has it been found necessary to take more than four terms to get a result correct to four significant figures.

12. It is not at first sight obvious that the determinant converges. Conditions sufficient for convergence are that the product of the diagonal terms should converge absolutely, and that the sum of the non-diagonal terms should converge absolutely.* The diagonal terms, except the first two, are all unity, and therefore the first condition is fulfilled.

It is clear that if we write

$$c_m' = \frac{2}{\pi^2}\left[\frac{b_m'}{m^2} \cdot \frac{\pi^2}{6} + \Sigma' \frac{b_n'}{n^2} \cdot \frac{1 + \cos(m-n)\pi}{2(m-n)^2}\right],$$

$$b_m' = \frac{4}{\pi^2} \Sigma' a_n \frac{1 - \cos(m-n)\pi}{2},$$

and call α_{mn}' the coefficient of a_n in c_m', then $|\alpha_{mn}| < \alpha_{mn}'$. If m is even,

$$c_m' = \frac{2}{\pi^2}\left[\frac{b_2'}{2^2}\frac{1}{(m-2)^2} + \frac{b_4'}{4^2}\frac{1}{(m-4)^2} + \dots + \frac{b_m'}{m^2}\frac{\pi^2}{6} + \dots\right],$$

---

* Whittaker and Watson, 'Modern Analysis,' § 2·81.

and α_{mn}' (n odd) is obtained from this series by replacing all the coefficients b by $4/\pi^2$; α_{mn}' is therefore independent of n. It is then clear that

$$\alpha_{2n}' + \alpha_{4n}' + \alpha_{6n}' + \ldots < \frac{8}{\pi^4}\left(\frac{1}{2^2} + \frac{1}{4^2} + \ldots\right)\left\{\frac{\pi^2}{6} + 2\left(\frac{1}{2^2} + \frac{1}{4^2} + \ldots\right)\right\} = \tfrac{1}{12}.$$

It can be seen in the same way that if n is even

$$\alpha_{1n}' + \alpha_{3n}' + \ldots < \tfrac{1}{4}.$$

Consequently $\sum\limits_{(m)} \alpha_{mn}'$ converges to a sum less than $\frac{1}{4}$ for all values of n.

It follows that the series $\sum\limits_{(m)} \alpha_{mn}$, and, *a fortiori*, the series $\Sigma\, \alpha_{mn}/m$ converge absolutely, and in no case can the sum of the moduli equal $\frac{1}{4}$. Consequently the double series

$$\Sigma\Sigma \frac{\alpha_{mn}}{m\mathrm{C}_n}$$

converges absolutely by comparison with the series $\Sigma \left(\frac{1}{{}_4\mathrm{C}_n}\right)$. This establishes the convergence of the determinant.

13. The preliminary work made it likely that the value of l^2 leading to a minimum value of AB would be between 1 and 2; AB has therefore been calculated for the values 1·0, 1·2, 1·4, 1·6, 1·8 and 2·0 of l^2. As a matter of convenience a new variable x, given by

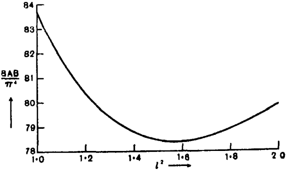
$$x = 8\mathrm{X}/\pi^4 = 8\mathrm{AB}l^2/\pi^4,$$

was used. The results are

l^2	1·0	1·2	1·4	1·6	1·8	2·0
x	83·7	96·4	110·3	125·5	142·1	159·8
$8\mathrm{AB}/\pi^4$	83·7	80·3	78·8	78·4	78·9	79·9

These are plotted in the figure and show that there is a minimum value

$$\mathrm{AB} = \frac{\pi^4}{8}(78 \cdot 4), \tag{40}$$



approximately, when l^2 is somewhat less than 1·6.

The results above were obtained from the first seven rows and columns of the determinant; two values of x differing by 5 and giving a change of sign were found, the seven-row determinants were calculated for these values, and the value of x which made the determinant vanish was calculated by linear interpretation.

The last figure given is not reliable, because of the method of interpolation, and also because of the slow convergence of the first two rows.

In the case $l^2 = 1\cdot6$ the determinant given by the first eight rows and columns was calculated.

When $x = 125$

$$\Delta_8 = (0\cdot0595)\,(1\cdot0025)\,(0\cdot2198)\,(0\cdot6392)\,(0\cdot6900),$$

while when $x = 130$

$$\Delta_8 = (-\,0\cdot0173)\,(0\cdot9904)\,(3\cdot9355)\,(1\cdot0761)\,(1\cdot0385);$$

the two values of Δ_7 are obtained by ignoring in each expression the last factor. Using the values of Δ_8, we find $x = 125\cdot4$; using Δ_7, $x = 125\cdot5$. It is clear that the stage has been reached at which the calculation of more rows and columns would lead only to unimportant arithmetical changes in the values of the determinants; the signs of the determinants evidently would not change. Since the first two rows converge slowly (like the series $\Sigma 1/n^2$), accurate numerical results would in any case be difficult to obtain.

14. From (40) the critical velocity for this type of disturbance can be calculated. Let N denote the Reynolds' number, $\bar{v}d/\nu$, of the undisturbed motion; $\bar{v}$ is the mean velocity across a section of the channel. Then

$$\mathrm{A} = 6\mathrm{N}/\pi^2, \quad \mathrm{B} = 12d\mathrm{N}/\pi^2 a,$$

and

$$8\mathrm{AB}/\pi^4 = 576\,\mathrm{N}^2 d/\pi^8 a.$$

The critical value of N is then given by

$$\mathrm{N} = 36\,(a/d)^{\frac{1}{2}}. \tag{41}$$

If then, for example, the distance between the cylinders is one-hundredth of the radius of the inner cylinder, so that $a/d = 100$, the critical velocity for this type of disturbance is given by $\mathrm{N} = 360$; if $a/d = 10$, $\mathrm{N} = 114$. The work of this paper will hardly apply to the second of these cases, wherein the value of d/a is probably too large; it has been assumed that d/a is small not only explicitly, but implicitly in ignoring the effect on the motion of the ends of the channel.

Some recent experimental work by S. J. Davies and C. M. White* shows that the lower critical velocity for flow through straight pipes of rectangular section is about 1400 when the width-depth ratio is large; and that provided this ratio is large enough (100 or over) its exact value is not important. Thus 1400 may be taken to be the critical velocity in a straight channel; compared with this value, those just calculated for a curved channel are small enough to suggest that (in a certain range of values of a/d) the type of disturbance considered here may be that which actually arises when stream-line motion breaks down. But this can hardly be decided otherwise than by experiment. an exploration of all possible small disturbances is impracticable.

15. It is perhaps necessary to show that the disturbance is damped out if N is less than the critical value, though this seems clear on physical grounds.

Equations (1) to (4) apply only to steady motions; if the motion is unsteady (4) is unchanged but $\frac{\partial U}{\partial t}$, $\frac{\partial V}{\partial t}$, $\frac{\partial W}{\partial t}$ must be added to the left sides of (1), (2), (3), respectively. If it is assumed that u, v, w, the components of the disturbed velocity, are proportional to $e^{-\sigma t}$ but do not otherwise depend on t, of the three fundamental equations (28) to (30), the last is unchanged, while the first two are

$$\left(\frac{\partial^2}{\partial y^2} - l'^2\right)\left(\frac{\partial w}{\partial y} + lu\right) = \mathrm{B}\left(\frac{y^2}{\pi^2} - \frac{y}{\pi}\right) lv,$$

$$\left(\frac{\partial^2}{\partial y^2} - l'^2\right) v = -\mathrm{A}\left(\frac{2y}{\pi} - 1\right) u,$$

where

$$l'^2 = \frac{d^2}{\pi^2}\left(\lambda^2 - \frac{\sigma}{\nu}\right) = l^2 - d^2\sigma/\pi^2\nu.$$

The equations in §§ 10, 11 are hardly altered, in some places, but not all, l' has to be substituted for l.

It is easy to see that the only changes needed in (37) are these. C_m must be replaced by C_m', where

$$C_m' = (l'^2 + m^2)^2/m,$$

and α_{mn} by α_{mn}', the latter being the same function of l' as the former is of l; lastly the $(m + 2)$th diagonal term is not now 1 but $(l^2 + m^2)/(l'^2 + m^2)$.

It is sufficient to take some numerical examples.

---

* 'Roy. Soc. Proc.,' A, vol. 119, p. 95 (1928).

If $l^2 = 1\cdot6$, (37) becomes

$$\left|\begin{array}{cccc} 1 & x\,(0\cdot01185) & 0 & \ldots \\ x\,(0\cdot00508) & 1 & x\,(-\,0\cdot00098) & \ldots \\ 0 & x\,(0\cdot00001) & 1 & \ldots \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \ldots \end{array}\right| = 0, \qquad (42)$$

after the first two rows and columns, which contain elements independent of x, have been eliminated. This is an equation to determine $x = 8\mathrm{AB}l^2/\pi^4$.

Suppose now that $l'^2 = 1\cdot6$, and that AB has a given value. We have to find the sign of σ, and this will be known when l is found. In this particular numerical case the equation to be solved is the same as (42) but for the diagonal terms. We have in fact

$$\left|\begin{array}{ccc} \dfrac{l^2+1}{l'^2+1} & x\,(0\cdot01185) & \cdot\cdot \\ x\,(0\cdot00508) & \dfrac{l^2+4}{l'^2+4} & \cdot\cdot \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot \end{array}\right| = 0 \qquad (43)$$

from which to determine l; x as before is $8\mathrm{AB}l^2/\pi^4$.

Equation (42) is satisfied if

$$x = 125\cdot5, \quad 8\mathrm{AB}/\pi^4 = 78\cdot4.$$

We have to show that for given l a value of AB less than the critical value leads to a positive value of σ.

Assume then in (43) that

$$8\mathrm{AB}/\pi^4 = 75,$$

for instance. If the diagonal terms in (43) were all unity, we should clearly have to take a value of l^2 greater than $1\cdot6$ to satisfy the equation; it would in fact be necessary to make

$$l^2/1\cdot6 = (78\cdot4)/75. \qquad (44)$$

But if $l^2 > l'^2$ the diagonal terms in (43) are greater than 1, and therefore the value of l^2 given by (44) will not make the determinant zero, but positive, and this requires that the value of l^2 should be still further increased. If l^2 is greater than l'^2, σ is positive, and the disturbance in this case is therefore damped out. This argument applies unchanged to any value of l'^2 greater than $1\cdot6$. There is a difference in the case of values of l'^2 less than $1\cdot6$. It is still clear that we must make $l^2 > l'^2$ to make the determinant vanish for a value of AB less than the critical for l'^2; but there is now the possibility that the value of AB taken may be greater than the critical value for l^2, for the critical value is a decreasing function for these values of l^2. The critical value of AB for l^2 is, of course, the significant value; the value for l'^2 is useful only as a guide to the arithmetic. In these cases, then, unlike the last, the two factors operating tend to cancel out. It therefore seemed necessary to treat one case numerically.

If $l^2 = 1 \cdot 2$, (37) becomes

$$\begin{vmatrix} 1 & x\,(0 \cdot 01536) & 0 & \cdots \\ x\,(0 \cdot 00662) & 1 & x\,(-0 \cdot 00156) & \cdot\,\cdot \\ 0 & x\,(-0 \cdot 00011) & 1 & \cdots \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot\,\cdot \end{vmatrix} = 0, \quad (45)$$

and this equation is satisfied by

$$x = 96 \cdot 4, \quad 8\mathrm{AB}/\pi^4 = 80 \cdot 3.$$

If $l'^2 = 1 \cdot 2$, we have an equation the same as (45) except for the diagonal terms in the determinant, which are the same as those in (43).

Taking $8\mathrm{AB}/\pi^4 = 80$, we find that

$$\Delta_5 = -214 \times 10^{-4}, \quad l^2 = 1 \cdot 23,$$
$$\Delta_5 = -108 \times 10^{-4}, \quad l^2 = 1 \cdot 22,$$

and can conclude that the determinant vanishes when l^2 is roughly $1 \cdot 21$. Since $l^2 > l'^2$ we have a positive value of σ, while from the figure it is clear that 80 is less than the critical value of $8\mathrm{AB}/\pi^4$ for $l^2 = 1 \cdot 21$.

16. An alternative method of showing that the disturbance is damped out unless the critical velocity is reached is to assume that the disturbed motion is steady as regards the time; and then to show that it decreases exponentially as it progresses along the channel. It would then be assumed that u, v, w were independent of t but were proportional to $e^{-\tau y}$, y denoting distance measured in the direction of flow along the centre of the channel. This appears to be a more satisfactory way of approach to the problem, the fluid must be supposed to enter the channel steadily but in a disturbed state; what has been found above is the velocity of flow for which $\tau = 0$ so that the initial state of disturbance persists along the channel. However, to prove that $\tau > 0$ unless the critical velocity is attained requires the solution of a problem algebraically very different from that of the special case $\tau = 0$,

and since it appears clear on physical grounds that this must be so, it has not been thought necessary to consider the matter in detail.

17. The work of this paper shows that flow in a curved channel may become unstable for small disturbances This is an effect of curvature: in a straight channel the persistence of a small disturbance of the type considered is impossible. This is offered as a tentative explanation of the known absence of a marked critical velocity of flow in a curved pipe.

In a straight pipe the passage through the critical velocity is accompanied by a sudden increase in the loss of head, but no such sudden change has ever been observed in a curved pipe. A possible explanation is that flow in a curved pipe (but not in a straight pipe) may become unstable for small disturbances, for instability so arising is unlikely to lead to other than a gradual change in the loss of head. The corresponding problem in two dimensions which has been worked out here lends some support to this suggestion.

An extension of the analysis to the three dimensional motion of fluid in a curved pipe would be difficult. There is one similarity between the two problems; stability in a curved channel has been shown to depend on the value of N^2d/a, while the stream-line motion in a curved pipe, and possibly therefore the stability of this motion, depends on N^2r/R,* where r is the radius of the pipe and R the radius of the circle in which it is coiled. However, the stream-line motion in a curved channel does not depend on N^2d/a; that is to say, there is, in this case, no scale effect, while in the curved pipe there is. This means that (unless the critical value of N^2r/R is far smaller than is likely) we must not assume the simple parabolic distribution of velocity to hold in stream-line motion in a curved pipe; and it is necessary that the stream-line motion should be simple if the stability problem is to be treated successfully on these lines, for the component velocities of the stream-line motion are coefficients in the differential equations of the disturbed flow.

* W. R. Dean, 'Phil. Mag.,' vol. 5, p 673 (1928)

Some Remarks concerning the Production and Absorption of Soft X-Rays and Secondary Electrons.

By Erik Rudberg, Nobel Institute Experimentalfältet, Sweden.

(Communicated by O W. Richardson, F.R S.—Received August 2, 1928)

Introduction.

In the present paper an attempt is made to estimate the efficiency of the production of photoelectrons from metals by soft X-rays When due consideration is paid to the absorption of radiation and electrons in the target the number of photoelectrons for one quantum of radiation absorbed comes out of the order unity. This result in connection with efficiency measurements on soft X-ray production, using the photoelectric method, shows that the yield of such radiation, when metals are bombarded with electrons, is extremely small. A very much larger portion of the energy of the bombarding electrons, possibly all the rest of it, reappears in smaller units as energy of secondary electrons. The question as to the probable origin of these electrons is discussed, it appears most likely that at least a great part of these are initially free conduction electrons of the substance.

Before the manuscript of this paper had been completed, a paper by Prof. O. W. Richardson* appeared dealing chiefly with the same problems. Some of the results presented here have already been arrived at by the latter author. Since, however, my own reasoning in deriving these results differs in several respects from that of Prof. Richardson's, I have decided to publish the present paper in its original form, in the hope that the different treatment would in itself present some points of interest.

Photoelectric Efficiency.

In a paper recently communicated to the Royal Society† a report has been given of an investigation concerning the velocity distribution among the photoelectrons produced by soft X-rays, which the writer has carried out in Prof. O. W. Richardson's laboratory at King's College, London, during the last two years. From the results obtained with several different conductors used for the solid photoelectric target, it was concluded that the bulk of the electrons emitted were to be regarded as the secondary emission set up in the substance by fast primary electrons, directly excited by the incident radiation quanta. Adopting this view and using a value of the absorption coefficient for the radiation employed, which had been obtained in some preliminary experiments with gold leaf, an attempt was made to estimate the total number of such primary electrons formed in the substance, corresponding to the measured, mainly secondary emission from the surface of the target. This calculation, which, apart from the admittedly rather idealising assumptions made with regard to the processes occurring also involved the use of absorption coefficients for the electrons derived from measurements by others, indicated that the number of primary electrons excited throughout the volume of the substance receiving any radiation would be greater than the secondary emission escaping from the surface by a factor of the order ten.

---

* 'Roy. Soc. Proc,' A, vol. 119, p. 531 (1928).

† 'Roy. Soc. Proc,' A, vol. 120, p. 385 (1928)

In discussing this result it was pointed out that little was known so far about the efficiency of the incident quanta in producing primary electrons. The evidence on this point in the ordinary X-ray field, and for much greater wave-lengths in the case of gases, was that one primary electron was ejected for each quantum absorbed. Again, the smallness of the photoelectric currents which can be drawn from solids, even when illuminated by very intense ultraviolet radiation, at least were not in conflict with the view, sometimes expressed, that only a fraction of the incident energy is spent in producing photoelectrons in this case. In the soft X-ray experiments referred to it therefore seemed possible that the number of radiation quanta falling on the target could be greater than the number of primary electrons produced, as derived from the measured saturation current using the calculated proportionality factor.

Since this paper was written I have come across some results published by other experimenters in this field, which, when connected with the inference drawn from the above-mentioned calculation of the total number of primaries, appear to lead to the conclusion that the number of quanta is of the same order of magnitude as the number of primary electrons excited in the case of these rays. In an experiment by F. Holweck,* who has published extensive measurements on soft X-rays using an ionisation chamber for the detection of the rays, this investigator compares the photoelectric current from a clean zinc surface and the total ionisation current produced in a gas by the same radiation. They are found to be in the ratio 1 : 50 for 160 volts' radiation and 1 : 85 for 280 volts. Now it seems to be well established that in the case of a gas the absorption of the incident radiation is practically entirely photoelectric, every absorption act giving rise to a free electron with an energy of the same order of magnitude as that of the absorbed quantum. It is true that if the latter is a little greater than the ionisation energy of some particular level of the parent molecule a considerable fraction of all absorption acts may result in the expulsion of electrons of comparatively low energy, and that in a few other cases the absorbed quantum will only suffice to shift an electron from a deeper level to one nearer to the surface of the parent atom, whence no free electron will result. But it is obvious, from the whole of the experimental evidence collected in work of this kind, that such effects do not occur to any such predominating extent as to alter the order of magnitude of the number of fast electrons produced for each quantum absorbed In order to arrive at an estimate of the number of radiation quanta entering the ionisation chamber, it is further necessary to know how many positive ions each of the fast electrons on an average will give rise to before it has spent all of its energy at the start. Such measurements have been the subject of special investigations, and as a result of these it is usually stated that the electron produces one ion for every 30 or 40 volts of its energy, averaged over a great number of ionisation processes. These conclusions have been drawn from measurements of the number of positive ions produced in air by electrons of a given velocity. Lenard and Becker* have published tables for this quantity ("totale Sekundarstrahlung, S_0") as a function of the velocity of the primary electron. From these one may take that the number of ions produced by a 250 volt-electron is about seven (the nearest values given are 6·0 for 230 volt- and 11·9 for 409 volt-electrons). It appears probable that Holweck has used some such data to arrive at the figure 10^{-7} for the yield of radiant energy corresponding to H120 volt-electrons impinging on a tungsten surface; the details of the calculation in this respect are not given.

---

* "De la Lumière aux Rayons X," p. 71, Paris (1927).

Now the mean energy of a quantum of radiation, although strongly filtered in Holweck's experiments, is likely to have been somewhat less than the bombarding voltage. Also there is some loss of energy, though usually small, by the primary electron in escaping from the parent atom, so that the conditions in the ionisation chamber, when the tube is run at 280 volts, would appear to correspond fairly closely to primary electrons with about 250 volts energy. The number of primaries in the gas, then, using Holweck's ratio 1 : 85, would come out about 12 times the number of electrons escaping from the zinc surface when illuminated by the same radiation. This value is of the same order (ten) as the ratio calculated by the writer, of the total number of primary electrons inside a solid conductor to the measured number of electrons escaping from the surface. In this case, furthermore, the evidence of the velocity analysis to which the emitted electrons were subjected led to the conclusion that the greater part of the primary electrons were produced by the K_a-radiation from carbon (275 volts). It is true that the photoelectric emission caused by soft X-rays depends to a certain extent on the substance of the emitter and the state of the surface; but the variations which may be found in this respect are not sufficient to change the order of magnitude. The result of this discussion is, therefore, that the number of primary electrons produced in gases is the same as the number produced in solid conductors by the same soft X-radiation. If, as appears most probable, every quantum absorbed gives rise to a photoelectron in the former case, the number of primary electrons must be equal to the number of quanta also in the case of solid conductors. It hardly needs be said that both statements refer to the order of magnitude.

---

* 'Handb. Exp.-Physik.,' vol. 14, p. 253, Leipzig (1927).

Efficiency of Soft X-Ray Production.

Reference has already been made to a statement by Holweck, who, using an ionisation chamber, finds a yield of radiant energy 10^{-7} with 120 volts across the X-ray tube in a typical case Making use of the results arrived at in the preceding paragraph it is possible to estimate the number of quanta emitted from the anode of the X-ray tube for each bombarding electron from measurements of the emission from a photoelectric target exposed to the radiation. Measurements of this kind have been published by Richardson and Chalklin,* and later by Richardson and Robertson.† Thus, for instance, Richardson and Chalklin give the following mean values of the experimentally determined ratio i_p/i_e (i_p photoelectric, i_e thermionic current): for carbon 1·57, for tungsten 6·29, for nickel 3·82, and for iron 4·20, all $\times 10^{-9}$. To get the order of magnitude we may take a figure somewhere in the middle of these, say, 3·5. The authors also state that the geometric factor by which these figures should be multiplied to obtain the ratio corresponding to the full radiation in all directions from the target is $4\pi/0{\cdot}0455 = 2{\cdot}76 \times 10^2$, which, using the mean value 3·5, makes the latter ratio come out $9{\cdot}5 \times 10^{-7}$. All this refers to bombarding electrons of 300 volts energy It is therefore possible that the mean wave-length was somewhat greater than in the experiments of the present writer, and the absorption coefficient for the rays correspondingly higher. From the deduction of the total number of primaries published in the writer's earlier paper, however, it appears unlikely that the change could be large enough to affect seriously the order of magnitude of the calculated proportionality factor; it should perhaps be expected to suffer a small decrease. If this factor is taken to fall between 1 and 10 it would seem to follow that the number of quanta emitted is in the ratio 10^{-6} or 10^{-5} to the number of bombarding electrons. Only one out of 10^5 or 10^6 electrons impinging on the anode is thus capable of producing a radiation quantum escaping from the latter This may be compared with the efficiency in the case of hard X-rays. Recent absolute measurements by Aurén,* using the total radiation from a tungsten anticathode at 100 *k*V., showed the energy falling on 1 cm ^{2} in a distance of 100 cm. to be 42 ergs per second when the tube was run at 1 milliamp. From the theoretical formula deduced by Kramers,† this figure should come out 50 This is equivalent to a yield of energy 5×10^{-3}, which is about a thousand times more than the yield in the soft X-ray region. Now according to Kramers' theory the efficiency should be proportional to the bombarding voltage in agreement with the empirical formula of Beatty.‡ This is for the continuous spectrum. In the case of soft X-rays it is not known yet how much of the radiation is continuous and how much characteristic; also the deductions of Kramers would certainly not apply to these rays without modification; but it will be seen that the order of magnitude is about what would be required by this formula for hard X-rays. On the other hand it should be remembered that the efficiency of excitation by electron impact in gases in the optical and ultraviolet region is quite considerable. in fact, in some of the cases which have been studied the yield appears to be of the order unity. In comparison with such a figure then the ratio of 10^{-6} or 10^{-5} here calculated, although it must be left undecided which of these comes nearer to the true value, is certainly surprisingly low. It would therefore appear to be necessary to look for some other process than the emission of radiation, by which the energy of the bombarding electrons is ultimately spent. It seems fairly certain that this process must be the production of a great number of slow-moving secondary electrons.

---

* 'Roy. Soc. Proc.' A, vol. 110, p. 247 (1926).

† 'Roy. Soc. Proc.' A, vol. 115, p. 280 (1927).

Soft X-Rays and Secondary Electrons.

Numerous investigations of recent years have established that there is an abundant emission of electrons from the surface of metals subjected to electron bombardment. The secondary current per unit bombarding current is of the order unity for a bombarding voltage of a couple of hundred volts, where this ratio appears to attain a very broad maximum. The velocity of the escaping secondary electrons shows a distribution rather similar to a Maxwellian one; the mean energy appears to be about 8 or 9 volts.* Since the electrons have to perform work against the opposing field in the surface, which is equivalent to some 4 or 5 volts in most cases, in getting out of the metal, the mean energy corresponds to about 13 volts for an electron inside the metal. If the bombarding primary electron of, say, 200 volts velocity were to spend its energy entirely in producing such 13-volt electrons there would result 15 of these in the target. This figure is considerably greater than the ratio of about 1 : 1 of secondary to primary current experimentally determined. In considering this point, however, several things have to be kept in mind. In the first place, absorption must always make the number of secondary electrons escaping from the surface smaller than the total number inside; the absorption coefficients for slow electrons are not much known at present, but might well be such as to account for the greater part of this difference. Secondly, it may be taken as certain, that of the escaping secondary electrons a considerable fraction has suffered loss of energy in collisions inside the target, which finally appears as Joule energy. This makes the mean energy of these electrons come out too low, when estimated from the observed velocity distribution in the emission from the surface. Thirdly, there are probably some secondaries the energy of which is insufficient to overcome the surface field; they remain in the target and are, therefore, not counted It is true that the energy of these electrons being small—less than 4 or 5 volts—their contribution to the energy balance would not be important, unless they were very numerous. I think that when all these circumstances are considered one must admit that nearly the whole energy of the bombarding electrons may be transferred into energy of secondary electrons inside the target. Of course, there is a third possibility, that of a primary electron giving up more or less of its initial energy directly to the surrounding atoms, thereby increasing the vibrational energy (heat motion) of the crystal lattice. That such a process must take place seems fairly certain, but it may well be that the less direct way by the production of slow secondary electrons is the one more frequently followed, both processes leading to the same resulting change. Since little is known about the mechanism of the former (direct) process, we shall disregard this alternative at present and consider the emission of secondary electrons, which process, possibly the dominating one, is certainly one very frequently occurring.

---

* Medd. K. Vet. Akad Nobel Inst., vol. 6, no. 13 (1925).

† 'Phil. Mag.,' vol. 46, p. 836 (1923).

‡ 'Roy. Soc. Proc.,' A, vol. 89, p. 314 (1914).

---

* This quantity is not to be obtained with any accuracy, since there is no sharp distinction between the secondary electrons proper and the primary electrons "reflected" from the surface retaining more or less of their initial energy. In some recent unpublished experiments, using a magnetic method similar to the one previously employed in the work on photoelectric emission, I found the mean energy for the total emission from copper with a bombarding voltage of 250 volts to be 34 volts. Subtracting the contribution of reflected primary electrons the mean energy came out 9·8 volts. The corresponding figures for the photoelectric emission produced by soft X-rays of 270 volts' energy, as given in the paper referred to above, were 15 and 6·8 volts respectively.

There is the question of the origin of the secondary electrons. Do they start from definite discrete energy levels of the parent atoms or are they identical with the free electrons assumed by most theories of electrical conductivity which are associated with a continuous spectrum of energy values? In the case of gases which also exhibit the phenomenon of secondary emission, it seems pretty certain that the electrons most frequently emitted are the most loosely bound ones, the valency electrons of the molecule.* Only when the energy of the bombarding electron is very little in excess of some critical potential of the latter, a considerable fraction of the secondary emission may originate in the corresponding deeper level.

That electrons are expelled from the more external levels in the case of gases is, of course, to be concluded from the considerable emission of radiation belonging to the visible and the ultraviolet region which generally accompanies the passage of electricity through gases. The radiation is emitted in the recovery process of the excited and ionised atoms which may take place directly or by successive steps. All these things are well known. The interesting point in this connection, however, is that there is nothing to correspond to this radiation in the case of solid conductors bombarded with electrons. Bricout,† using a powerful spectrograph, was able to get a blackening of a photographic plate in the region 2100 to 2300 Å.U. in a 20-hours' exposure to the radiation from a platinum anode bombarded with one milliamp. under 6 volts, he believes that this radiation could not have been due to residual gas. In the case of 700 volts' bombarding electrons and a carbon anode I have found that the insertion of a fluorite window in the beam of emergent radiation cuts down the photoelectric emission from copper in the ratio 600 : 1; the remaining current in this case might well be due to gas radiation. In both cases, even if the radiation is actually characteristic of the solid conductor itself, the amount is vanishingly small compared with the amount radiated by gases under similar conditions. The only case which I am aware of, where a luminosity other than such due to high temperature has been observed with solid conductors, is in the work on pulling electrons out of metals by Millikan and Eyring.* These authors sometimes noticed the occurrence of luminous spots on the cylindrical anode surrounding the narrow central filament when the current was increased sufficiently. The spectrum appeared to be continuous. Since, however, the current in this case was believed to come from extremely minute irregularities of the filament surface opposite the bright spots, the bombardment was presumably very intense over the small area of the latter; also the voltage was rather high, 4,000–12,000 volts. It seems possible, therefore, that some extraordinary conditions—local vaporisation, patches of high temperature (continuous spectrum)—were responsible for the luminosity in this case. For electrons of moderate speed—a few hundred volts—bombarding solid conductors it appears rather certain that there is nothing to correspond to the considerable radiation from gases. Among the latter it may be noticed that metal vapours are generally distinguished for high luminosity.

---

* Franck and Jordan, "Anregung von Quantensprüngen durch Stösse," p. 184, Berlin (1926).

† 'C. R.,' vol. 182, p. 213 (1926).

This difference should probably be regarded as a fundamental feature of the processes occurring in solid conductors subjected to electron bombardment. We shall now return to the question of the origin of the secondary electrons. On the first alternative, that of electrons originally bound and having discrete possible energy values, it will be necessary to account for the absence of any radiation when the ejected electron is replaced by a new one. In the first place one might try the hypothesis that there is an emission of radiation, but that this is so completely reabsorbed in the substance that the fraction which is able to escape from the target is too small to be detectible. However, experiments† have shown that the absorption coefficients for the slow secondary electrons and radiation of the corresponding frequency are of about the same order of magnitude To get the enormous absorption required on this hypothesis it would therefore seem necessary to assume that reabsorption takes place with an abnormally high probability within the parent atom itself or in the nearest atoms, so to speak, before the radiation has settled down to the state in which it is found in the usual kind of absorption experiments This would be something very similar to the compound photoelectric effect studied by Auger‡ in the case of ordinary X-rays only that the probability of reabsorption in the present case would have to be much higher than any of the values found by Auger. It is perhaps doubtful if processes of this character should be described as production and reabsorption of radiation at all; the effect would at any rate be very much the same as the one which might next be tried in order to account for the absence of radiation from atoms contributing their bound electrons to the secondary emission on the first alternative: rearrangement of the disturbed orbits without the emission of any radiation. This is no doubt an effect which actually does occur in many cases. The quenching of fluorescence in gases produced by increased pressure or by the admixture of foreign gases is evidently of this kind; in the latter case the highly specific quenching by different gases rather indicates that there is some essential condition of coupling or resonance of the atomic systems involved which may possibly be more conveniently described as emission and reabsorption of radiation. Such "radiationless" transformations are thus by no means uncommon among the low energy levels here considered, and they might well be expected to play an important part where the atoms are as closely packed as in solids.

---

* 'Phys. Rev.,' vol. 27, p. 51 (1926).

† Partzsch and Hallwachs, 'Ann. Physik.,' vol. 41, p. 247 (1913).

‡ 'Ann. Physique,' vol. 6, p. 183 (1926)

If, on the other hand, the secondary electrons are all initially free ones it is fairly easy to account for the absence of radiation in experiments on electron bombardment. The emitted electrons are simply replaced by other free electrons, the energy difference—positive or negative being of the same order as the amounts perpetually exchanged between the free electrons and the surrounding atoms in thermal equilibrium. Obviously any radiation emitted in this case would be indistinguishable from the pure temperature emission from the target.

Thus far there appears to be nothing to decide between the two hypotheses as to the origin of the secondary electrons, although the last one would seem the more simple of the two. There is, however, one experimental fact which appears to me to indicate that the free electrons play an important part in the processes actually occurring It is known that a very great number of substances other than gases become emitters of visible radiation under electron bombardment. With liquids this fluorescence generally ceases immediately the bombardment is discontinued, whilst many solids exhibit the phenomenon of phosphorescence. Both in the case of liquids and solids showing luminescence the packing is generally almost as close as in metals. Further, it would appear that luminescence under electron bombardment is a rather common phenomenon which is found with a great number of different substances. It is possible that this property of matter is still a good deal more general than hitherto known, since the radiation emitted in some cases could fall outside the visible part of the spectrum. It seems to me a point of some significance that in no case, so far as I am aware, has luminosity under electron bombardment been found with a substance showing metallic conduction. It would therefore appear very likely that the absence of radiation in the case of solid conductors is intimately connected with the presence of free electrons in such substances.

This argument, however, is not necessarily against a mechanism involving the interaction of excited atoms. From thermodynamical reasoning Klein and Rosseland* have concluded that slow-moving electrons should be very efficient in suppressing the emission from excited atoms owing to the frequent occurrence of so-called collisions of the second kind as a result of which the initially slow electrons acquire the energy of excitation of the atoms hit as kinetic energy, leaving the atoms in the normal state. The concentration of slow electrons in a good conductor being enormous compared with the electron concentrations which it is possible to use in experiments on gases, it may well be that this process is responsible for the extinction of practically all excited atoms which otherwise would have emitted radiation.

It is not necessary that this process should be confined to the case of the most external levels of the atom. It seems possible that an effect of this kind could be the cause of the low efficiency of soft X-ray production in general referred to above. The result of this effect would again be the same as if radiation were emitted, but immediately absorbed giving rise to a photoelectron, but the present explanation avoids the introduction of abnormally high absorption coefficients for the radiation In both cases electrons of the corresponding energy, less than that of the bombarding electrons, should result, and ought to show up in the velocity spectrum of the emission from the surface, if such effects occur to any considerable extent. So far no sign of such electrons has been found in the case of the heavy metals that have been tested in this respect.† It is possible, however, that the search would be more likely to be successful if a comparatively simple substance of low atomic number and with a fairly well-defined critical potential of the order of a few hundred volts, such as carbon, could be studied. The writer is hoping shortly to be able to try this experiment.

Returning to the low energy secondary emission it would appear that one of the most serious difficulties at present for any theory which postulates a mechanism involving discrete energy levels, is to account for the total absence of any structure in the low velocity spectrum of the secondary electrons escaping from the surface. It seems rather improbable that subsequent energy losses of the secondary electrons on their way to the surface could entirely destroy every sign of an upper limit of velocity corresponding to the initial energy of the excited atoms responsible for the emission. Also the great similarity of the distribution curves for different elements hardly suggests any intervention of energy levels, characteristic of the element in each particular case. Indeed, photoelectric experiments in the visible and ultraviolet region have established that the velocities vary in a marked and regular manner with the frequency of the incident light. On the other hand, the occurrence of rapid increases in the total secondary emission at certain critical potentials, as the bombarding voltage is increased, would seem to indicate that a new energy level is excited as the particular voltage is passed. In the case of tungsten, Richardson and Chalklin* have recently shown that there is a beautiful agreement between their own very carefully determined critical potentials for soft X-ray excitation and the very accurate values published by Krefft† for the critical potentials of secondary emission. Farnsworth,‡ however, has found that the position and appearance of many of the discontinuities in the secondary emission curves depend on the crystal structure of the target and of the orientation of the latter in the case of single crystals He is, therefore, inclined to regard the peculiarities in the curves as resulting from selective reflections occurring in the crystal lattice It is to be hoped that a detailed study of the soft X-ray emission from solids by grating methods, which now seems to be possible, will definitely settle the question as to the true significance of most critical potentials.

---

* 'Z. Physik,' vol. 4, p. 46 (1921).

† It is possible, however, that the anomalous velocity distribution found for the electrons transmitted through thin metal foils in some experiments by A. Becker ('Ann. d. Physik, vol. 84, p. 779 (1927)) is due to some process of this kind.

It would follow from this discussion that whether the secondary electrons appear as a consequence of a preceding excitation of the atoms in the conductor, as on the first alternative, or are produced more directly by the primary electrons as on the second, they would largely consist of initially free electrons inside the conductor In the latter case the emission would be made up entirely of such electrons. So far the present considerations have been confined to the question of the origin of the secondary electrons. In particular, no suggestion has been made as to the nature of the process by which part of the energy of the primary electron would be transferred to a free electron of the conductor on the second alternative. In a recent interesting paper by Prof. Richardson, referred to in the beginning of the present note, a theory is developed according to which this transfer involves the interaction of radiation. It is also announced that the adoption of this view will remove the most obstinate difficulties of the auto-photoelectric theory of thermionic emission and chemical reactions in general. Further applications will probably show the fruitfulness of this conception.

---

* 'Roy. Soc. Proc.' A, vol. 119, p. 60 (1928).

† 'Ann. Physik,' vol. 84, p. 639 (1927), and 'Phys. Rev.,' vol. 31, p. 199 (1928).

‡ 'Phys. Rev.,' vol. 31, p. 419 (1928).

Hyperfine Structure in the Arc Spectrum of Cæsium and Nuclear Rotation.

By D. A. Jackson, Clarendon Laboratory, Oxford.

(Communicated by F. A. Lindemann, F.R.S.—Received August 3—Revised September 18, 1928.)

Summary.

The arc spectrum of cæsium was investigated with the object of finding whether any of its lines possessed hyperfine structure, resulting from a nuclear magnetic moment, due to a quantised nuclear spin. The lines belonging to the principal series should, owing to the greater degree of penetration of the electron in the ($1s$ or 6_1) orbit, and the correspondingly greater interaction, show the greatest effect. The lines of the principal series are very easily broadened if the vapour pressure of the metal becomes high, so that great care had to be used in obtaining the spectrum of cæsium at a sufficiently low temperature. The most satisfactory method of excitation was found to be the application by means of external electrodes of a very high frequency alternating current to a tube filled with helium at about 2 mm. pressure containing a small quantity of cæsium. The tube required slight heating to bring out the cæsium lines; without this the helium spectrum was very much stronger than the metallic spectrum. At a very low vapour pressures of cæsium the discharge was blue in colour. Under these conditions the lines of the principal series showed no broadening greater than that due to thermal agitation, but at a slightly higher temperature the colour of the discharge became purple and the lines broadened. The lines belonging to the principal series were found to be very close doublets

with very nearly constant frequencies differences. A theory is worked out which explains the origin of these doublets, assuming a nuclear spin of one half quantum; by correlating the difference in the separation of the hyperfine structure doublets in the $1s - m\,{}^2p_{3/2}$ lines and the $1s - m\,{}^2p_{1/2}$ lines, it is shown that a ratio of the magnetic to the mechanical moment of the nucleus about twice as great as the corresponding ratio for the electron would account for the observed frequency differences. The spectral notation used throughout is that of Hund.*

The results are compared with those found for the hyperfine structure of some of the bismuth lines by Back and Goudsmid,† and are found to be in satisfactory agreement. A selection principle is found which applies both to the bismuth and the cæsium spectrum.

Method of obtaining the Spectrum of Cæsium.

A tube made of Pyrex glass with the capillary portion about 5 cm. long and of about 8 mm. internal diameter, and the ends about 7 cm. long and 3 cm. in diameter, was filled with helium carefully purified by charcoal immersed in liquid air to a pressure of about 2 mm. The cæsium was distilled into the tube through four bulbs, each one being sealed off as the metal was distilled from it, into the spectrum tube.

The tube was excited by connecting it by means of large copper sheeting external electrodes to an oscillator of frequency of about 10,000 kilocycles. The oscillating circuit used was that of Gutton.‡ Two ½-kw. transmitting valves were used, the plate potential difference being about 2000 volts and the plate current about 0·5 amp. Under these conditions the tube was very strongly excited when one of the external electrodes was connected, to one end of the plate coil of the Gutton oscillator. When the tube was cold the helium spectrum was produced with very great brilliance, together with a few weak cæsium lines.

Slight warming of the tube rendered the discharge blue in colour and extinguished the helium spectrum completely. The presence of drops of metal in the capillary portion of the tube had the same effect. Most of the visible

---

* *Notation of spectral terms.*—

Fowler	$m\sigma$	$m\pi_2$	$m\pi_1$	$m\delta_2$	$m\delta_1$
Sommerfeld	ms	mp_1	mp_2	md_1	md_2
Hund (used here)......	ms	$m^2p_{1/2}$	$m^2p_{3/2}$	$m^2d_{3/2}$	$m^2d_{5/2}$

† 'Zeit fur Phys.,' vol. 43, p. 321 (1927), and vol. 47, p. 174 (1928).

‡ 'Onde Electrique,' vol. 4, p. 387 (1925).

light came from the $1s - 3p$ lines, under these conditions. More intense heating of the tube made the discharge become reddish purple, the lines of the diffuse series, which are situated mostly in the red and infra-red becoming stronger. At the same time the lines of the principal series became very much broadened, particularly the first members.

Still greater heating changed the colour of the discharge to golden yellow. This corresponded to the development of the spark lines of cæsium, which are situated in the yellow and green. A. Filipov and E. Gross* measured the fine structure of some of these spark lines, and remark that all the arc lines of cæsium are simple. But if the vapour pressure of the cæsium is sufficiently great to give rise to the spark spectrum, as was the case at the temperatures with which Filipov and Gross worked, the arc lines belonging to the principal series are very seriously blurred and the hyperfine structure is completely obscured. Accordingly, for examination of the structure of these lines the tube was used at the temperature at which the discharge just appeared blue.

Measurement of the Hyperfine Structure of the Lines.

The spectrum was photographed with a 3-metre spectrograph of the Littrow type used in conjunction with Fabry and Perot etalons. In the blue and ultra-violet a quartz lens and prism were used, and in the red and infra-red a dense glass prism and lens. For photographing the second and third members of the principal series, Ilford, Iso, Zenith, and Wellington anti-screen plates were used with equal success. The lines in the near red were photographed on Ilford Rapid Process Panchromatic plates. In the infra-red down to 7500 Å.U., plates stained with dicyanine A were used. Wellington anti-screen plates were found to be the most satisfactory for this purpose. Five milligrams of the dye were dissolved in 60 c.c. of water, 40 c.c. of methyl alcohol, the solution also contained 5 per cent. by volume of ammonia solution of density 0·88. In the extreme infra-red Neocyanine plates, supplied by Eastman's, were used. These were also hypersensitised with 5 per cent. ammonia solution. In both cases the plates were bathed for one minute, and dried rapidly in front of an electric fan. The etalons used were made by Adam Hilger, the separating rings being of fused silica and the plates of quartz. The ring system was projected on to the spectroscope slit in the customary manner. The condenser was placed so as to collimate the light from the cæsium tube; in this way the maximum illumination is obtained; also the image of the tube is accurately focussed on the slit, so that the arrangement is completely stigmatic. The condenser was of about 30 cm. focal length. The etalon objective used in the blue and ultra-violet was a quartz fluorite achromat of 15 cm. focal length, in the red and infra-red a crown glass lens of 10 cm. focal length was used. The etalons employed were 2·5, 5, 10 and 20 cm long. The fringes were focussed on to the slit first by visual observation approximately. Then a trial plate was taken in which the etalon objective was moved by successive millimetres over a range of 5 mm. on each side of the approximately observed visual reading. The best defined of these was then selected.

---

* 'Z. Physik,' vol. 42, p. 77 (1927). In a later paper (*tom. cit.*, p. 497), Filipov published photographs of the $1s - 3p$ lines taken at lower temperatures with an echelon grating. The lines can here be seen to be double: owing, however, to the small dispersion Filipov mistook the structure for self-reversal. It is interesting here to note that the present author was able in the $1s - 3p$ lines to observe the self-reversal of each of the hyperfine components of these lines, by suitable adjustment of the temperature.

The lines of the principal series were found to be close doublets of equal intensity just resolved by the 2·5 etalon; the 5 mm. etalon showed them to be in all cases about 0·3 fringe apart and the 10 mm. etalon showed them to be about 0·6 fringe apart. In the $1s - 2\,{}^2p_{3/2}$ line each of the components could be obtained self-reversed. Many photographs were taken of all of the first three doublets of the principal series. The fringes were then measured on a Hilger photomeasuring micrometer. As is well known the square of the diameters of the rings varies linearly with the order of interference (P). The first four or five rings were measured. The diameters of these were then squared and from these figures the separation of the doublets was calculated. Examples of such measurements are given in order to demonstrate the accuracy of the measurements.

The results of these measurements are summarised in Table I. Each figure represents the mean separation of the four or five ring diameters measured in each fringe system. In addition to these lines the fringes of all the stronger lines of the diffuse and sharp series were photographed. They were photographed together with the standard Neon lines* with the 5 mm. etalon so that values of their wave-lengths accurate to about 0·001 Å.U. were obtained by combining the integral number of the order of interference, as derived from the measurements of Meissner,† with the exact fractional values derived from the measurements of the fringe systems. These figures are, however, not given here as they are scarcely within the scope of the research; moreover the values of Meissner may be in error to a greater extent than 0·1 Å.U., the accuracy required to fix the integral number of the order of interference with a

---

* 'Bur. Standards Bull.' vol. 14, p. 769 (19).

† 'Ann. Physik,' vol. 65, p. 378 (1921).

5 mm. etalon; in this case the values given would be meaningless. A list is appended of those lines whose fringes were photographed and found to be simple. Among these are the lines $2\,{}^2p_{3/2} - 3s, 4s, 5s$ and $2\,{}^2p_{1/2} - 3s, 4s, 5s$. It might at first be thought that these should be doublets, but a simple calculation shows that the separation is so small, even in the case of the $2p - 3s$ lines, that it is smaller than the natural breadth of the lines due to the Doppler effect of the gas kinetic velocities of the cæsium atoms, and therefore unobservable.

Accuracy of Measurement of Frequency Differences of Principal Series Doublets.

The table of measurements gives a fair idea of the consistency of the observations. The measurements for the $1s - 4\,{}^2p_{1/2}$ line were more uncertain owing to the fact that the very strong helium line at 3888 lies right over this line; the tube had therefore to be kept sufficiently hot to extinguish the helium lines completely; for the helium line, when the tube was cold, required only about 3 seconds exposure, while the cæsium line needed about $\frac{3}{4}$ hour.

It can be observed that the probable error due to the accidental errors varies from less than 0·005 for the lines $1s - 2p$ and $1s - 3p$ to about 0·010 for the $1s - 4p$ lines. The probable error is taken as the mean deviation of any individual figure divided by the square root of the number of determinations. In addition to these erratic errors, there may be also systematic errors of about the same order of magnitude. These arise from the circumstance that the resolving power of the etalon is not large compared with the separation of the lines. It varies from about 0·2 to 0·1 of a fringe, according to whether it is being used in the region of short or long wave-lengths. This resolving power with the 10 mm. etalon gives the line a breadth rather larger than the natural breadth of the lines arising from the blurring due to the Doppler effect. In the case of adjacent $1s - m\,{}^2p_{3/2}$ and $1s - m\,{}^2p_{1/2}$ lines the systematic error is very nearly equal in magnitude, for the resolving power of the etalon is almost the same for two such lines. But in the case of lines so widely separated as the $1s - 2p$, the $1s - 3p$ and the $1s - 4p$ doublets, which vary in wave-length from 9000 Å.U. to 3800 Å.U., the resolving power of the etalon varies from about 10P to about 4P, owing to the better reflecting power of the silver surfaces for the light of longer wave-length (P = order of interference). The systematic errors will not therefore be of equal magnitude for these widely separated lines. The origin of these systematic errors is the disturbing effect which the proximity of a neighbouring line has on the apparent position of the line as measured on the micrometer.

The upper limit of the magnitude of the systematic error is derived from a comparison of the values of differences in order of interference as measured with the 10 mm. and the 5 mm. etalons. The error in the case of the 5 mm. etalon is very much greater than the corresponding error for the 10 mm. etalon. For the distance between lines of the same order is about 0·3 of a fringe; with the 10-mm. etalon the corresponding distance is 0·6 of a fringe; and the separation of lines of order of interference differing by 1 is 0·4 fringe. The systematic error is therefore smaller with the longer etalon, the lines being more widely separated and consequently causing a smaller disturbing influence in the measurements. Moreover the error is in the opposite direction in the two cases; for the effect of the proximity of two lines is to cause the observer to over-estimate the distance between them. Thus with the 5 mm. etalon the apparent separation of the lines is greater than the true separation, but with the 10 mm. etalon the distance of 0·4 of a fringe, that between lines of adjacent orders of interference will be overestimated, and the separation of 0·6 of a fringe of the doublet will be consequently underestimated.

The systematic error is probably about twice as great with the shorter etalon, as is also the erratic error. This being assumed to be the case, it can be seen that a correction of about 0·008 fringe subtracted from the 5 mm. etalon reading, and a correction of about 0·004 added to the 10 mm. reading, reduces the measurements of all the fringes to agreement. This correction is not actually made but its magnitude is given for each line. The correction is about 0·004 for $1s - 2p$ and $1s - 2p_2$, $1s - 3p_1$ and $1s - 3p_2$ and $1s - 4p_2$; while for $1s - 4p_1$ it appears to be very nearly zero; this effect is, however, obviously a fortuitous one and well within the probable erratic error; the systematic error can therefore be said to be not greater than 0·005 for all the doublets measured.

From the table of measurements it is evident that the separation of the doublets is almost constant when measured in differences in order of interference; as is well known this corresponds to constant frequency or wave number differences; for this reason the results are converted into wave numbers and not wave-lengths. This is also advantageous *a priori* from the theoretical point of view. The frequency differences are expressed in cm.$^{-1}$; these are obtained by dividing the difference measured in terms of fractions of a fringe, by twice the thickness of the etalon.

Table of Measurements.

Each of the figures represents the results of the measurement of one fringe system; that is, it is given by the mean value derived from the separations of four or five double rings.

Differences in Order of Interference with 5 mm. Etalon.

$1s - 2\,{}^2p_{1/2}$.	$1s - 2\,{}^2p_{3/2}$.	$1s - 3\,{}^2p_{1/2}$	$1s - 3\,{}^2p_{3/2}$.	$1s - 4\,{}^2p_{1/2}$.	$1s - 4\,{}^2p_{3/2}$.
0·320 ±0·015 0·330 ±0·015	0·300 ±0·010 0·295 ±0·010	0·322 ±0·025 0·317 ±0·020 0·315 ±0·010	0·313 ±0·020 0·309 ±0·010 0·309 ±0·010	0·312 ±0·020 0·292 ±0·020 0·298 ±0·020	0·307 ±0·015
0·325 ±0·010	0·298 ±0·007	0·317 ±0·010	0·310 ±0·010	0·309 ±0·015	0·307 ±0·015

Differences in Order of Interference with 10 mm. Etalon.

$1s - 2\,{}^2p_{1/2}$.	$1s - 2\,{}^2p_{3/2}$.	$1s - 3\,{}^2p_{1/2}$.	$1s - 3\,{}^2p_{3/2}$.	$1s - 4\,{}^2p_{1/2}$.	$1s - 4\,{}^2p_{3/2}$.
0·620 ±0·010 0·630 ±0·010 0·622 ±0·010 0·633 ±0·010 0·624 ±0·010 0·630 ±0·010	0·588 ±0·004 0·587 ±0·004 0·589 ±0·004 0·588 ±0·004	0·613 ±0·010 0·622 ±0·010 0·612 ±0·010 0·620 ±0·010 0·616 ±0·010	0·599 ±0·010 0·605 ±0·010 0·600 ±0·010 0·600 ±0·015 0·595 ±0·010	0·600 ±0·020 0·615 ±0·010 0·619 ±0·020 0·608 ±0·015	0·600 ±0·020 0·600 ±0·015 0·590 ±0·010 0·610 ±0·020 0·595 ±0·010
0·629 ±0·005	0·588 ±0·002	0·617 ±0·005	0·600 ±0·005	0·611 ±0·010	0·598 ±0·010

Frequency Differences in cm.$^{-1}$. (From 10 mm. Etalon.)

$1s - 2\,{}^2p_{1/2}$.	$1s - 2\,{}^2p_{3/2}$.	$1s - 3\,{}^2p_{1/2}$.	$1s - 3\,{}^2p_{3/2}$,	$1s - 4\,{}^2p_{1/2}$.	$1s - 4\,{}^2p_{3/2}$.
0·315 ±0·003	0·294 ±0·001	0·308 ±0·003	0·300 ±0·003	0·305 ±0·010	0·299 ±0·005

Example of the Measurements of a Fringe System.

Micrometer readings.		Diameters.	Diameter².	Separations.
59·323	63·680	4 357	18·98	2·65
59·482	63·523	4 041	16·33	
59·590	63·420	3·830	14·67	2·65
59·768	63·235	3 467	12 02	
59·900	63·107	3 207	10·28	2 62
60·117	62·885	2 768	7·66	
60·277	62·723	2 466	6 08	2·72
60 580	62·412	1·832	3·36	

Mean separation = 2·66 ± 0 05.

Diameter² of three rings = 18·98 − 6·08 = 12 90
16·33 − 3 36 = 12·97,

therefore 1 diameter² = 4·31.

Therefore difference = 2·66 ± 0·05/4·31 = 0·616 ± 0·010.

Table of Lines Measured and Found Simple.

Line.	Wave-length.	Etalons used.
	Å.U.	
$2\,^2p_{3/2} - 3s$	7943	2 5 mm., 5 mm., 10 mm.
$2\,^4p_{1/2} - 3s$	7608	2·5 mm., 5 mm., 10 mm.
$2\,^2p_{3/2} - 4s$	6586	2·5 mm., 20 mm.
$2\,^2p_{1/2} - 4s$	6354	2·5 mm., 20 mm.
$2\,^2p_{3/2} - 5s$	6034	2 5 mm., 20 mm.
$2\,^2p_{1/2} - 5s$	5839	2·5 mm., 20 mm.
$2\,^2p_{3/2} - 4\,^2d_{3/2}$	8761	2·5 mm., 5 mm., 10 mm.
$2\,^2p_{3/2} - 5\,^2d_{3/2}$	6983	2·5 mm , 5 mm., 20 mm.
$2\,^2p_{3/2} - 5\,^2d_{5/2}$	6973	2·5 mm., 5 mm., 20 mm.
$2\,^2p_{1/2} - 3\,^2d_{3/2}$	6723	2·5 mm., 5 mm., 20 mm.
$2\,^2p_{3/2} - 6\,^2d_{3/2}$	6217	2·5 mm., 20 mm.
$2\,^2p_{3/2} - 6\,^2d_{5/2}$	6212	2·5 mm., 20 mm.
$2\,^2p_{1/2} - 6\,^2d_{3/2}$	6010	2 5 mm., 20 mm.
$2\,^2p_{1/2} - 7\,^2d_{3/2}$	5847	2·5 mm., 20 mm.
$2\,^2p_{1/2} - 7\,^2d_{5/2}$	5844	2·5 mm., 20 mm.

Interpretation of Results.

The investigation of the spectrum showed that the lines belonging to the principal series consist of close doublets, the separations for the first three members being in cm.$^{-1}$.

$1s - 2\,{}^2p_{3/2}$	$0\cdot294 \pm 0\cdot001$	$1s - 3\,{}^2p_{3/2}$	$0\cdot300 \pm 0\cdot003$
$1s - 2\,{}^2p_{1/2}$	$0\cdot315 \pm 0\cdot003$	$1s - 3\,{}^2p_{1/2}$	$0\cdot308 \pm 0\cdot003$
Differences	$0\cdot021 \pm 0\cdot004$		$0\cdot008 \pm 0\cdot006$

$1s - 4\,{}^2p_{3/2}$	$0\cdot299 \pm 0\cdot005$
$1s - 4\,{}^2p_{1/2}$	$0\cdot305 \pm 0\cdot005$
Difference	$0\cdot00_6 \pm 0\cdot01$

Each of these frequency differences may be regarded as composed of two parts; a separation of about 0·30 cm.$^{-1}$, which is independent of the line observed; and a much smaller part varying in amount with the current quantum number of the p term of the line to be added to the large separation for the $m\,{}^2p_{1/2}$ lines and subtracted for the $m\,{}^2p_{3/2}$ lines. The value of this part of the separations is given by the difference in the separation of the $1s - m\,{}^2p_{1/2}$ lines and the $1s - m\,{}^2p_{3/2}$ lines (fig. 1) represents a system of energy levels which would give the observed separations. The diagram is drawn only for the first doublet of the series; it would be the same for the lines $1s - 3p$ and $1s - 4p$, except that the differences in the p levels would be 0·004 and $0\cdot00_2$ cm.$^{-1}$ instead of 0·01 cm.$^{-1}$.

The arrows drawn in full represent the observed transitions and those dotted the unobserved transitions. In the diagram on the right the levels between which transitions are not observed are shown separated. It appears from this point of view that the cæsium atom possesses two sets of levels between which transitions either do not occur or occur much less frequently. In the set of levels identified by the letter α, the levels of the $s_{1/2}$ and $p_{3/2}$ terms are lowered and those of the $p_{1/2}$ term raised; while in the other set the reverse occurs.

The hyperfine structure of cæsium lines is thus represented empirically by the above energy diagram. This can be explained by attributing to the nucleus a rotation of one-half quantum; if this is done a result is obtained which agrees very satisfactorily with the analysis of the hyperfine structure of the bismuth lines made by Back and Goudsmid.*

The absence, or weakness of the intercombinations between the α levels and the β levels can be explained by the assumption that in the α levels the

---

* 'Z. Physik,' vol. 43, p. 321 (1927), and vol. 47, p. 174 (1928).

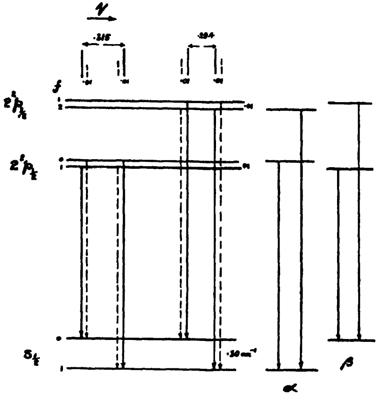
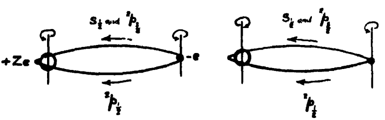


FIG. 1.

direction of the nuclear and electronic spins are the same, and in the β levels opposite, as shown in fig. 2.



	α			β	
	$\vec{j} + \vec{i} = \vec{f}$			$\vec{j} + \vec{i} = \vec{f}$	
$s_{1/2}$	$\frac{1}{2} + \frac{1}{2} = 1$	lower.	$s_{1/2}$	$\frac{1}{2} - \frac{1}{2} = 0$	higher.
$^2p_{3/2}$	$\frac{1}{2} + \frac{1}{2} = 0$	lower.	$^2p_{3/2}$	$\frac{1}{2} - \frac{1}{2} = 1$	higher.
$^2p_{1/2}$	$\frac{1}{2} - \frac{1}{2} = 0$	higher.	$^2p_{1/2}$	$\frac{1}{2} + \frac{1}{2} = 1$	lower.

FIG. 2.

The quantum numbers f represent the vectorial sum of the quantum numbers j and i ($j = l \pm s$; $s = \frac{1}{2}$; and i is the quantum of the nucleus) in the same manner as that adopted by Back and Goudsmid.

In the terms corresponding to the configuration α, j and i must be added for the $s_{1/2}$ and ${}^2p_{3/2}$ terms and subtracted for the ${}^2p_{1/2}$ terms; for in the first two, the direction of j is the same as that of electron spin, while in the last it is in the opposite direction. In the configuration β, j and i must clearly be subtracted for the s and ${}^2p_{3/2}$ terms and added for the $p_{1/2}$ terms. The reason for the choice of the number $\frac{1}{2}$ for the nuclear spin is that this value combines most simply with the half integral values of j so as to give a doublet structure to the levels. It is also clear that the higher value of f must correspond to the lower value of the energy level, in direct contradiction to the behaviour of the electron spin; the charge of the nucleus being positive and that of the electron negative.

The selection principle which require the intercombinations between the two configurations to be either absent or very weak is of great interest; particularly as the intensities in the bismuth lines can be shown to be in agreement with it. The accompanying diagram is taken direct from Back and Goudsmid's paper. The value of i is here $4\frac{1}{2}$ giving rise to the values of f shown in fig. 3. Here the two strongest lines are $f = 5$ to $f = 6$ and $f = 4$ to

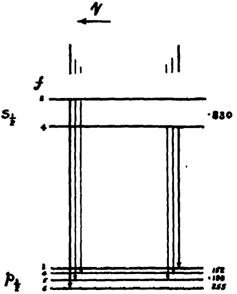


FIG. 3.—Fine Structure of Bismuth Line at 4722, according to Back and Goudsmid.

$f = 3$; now the first of these corresponds to the transitions where electron spin and nuclear spin are parallel in both final and initial states, there being no change in orientation during the transition; and the second corresponds to the electron and nuclear spins being antiparallel in both states, with here also no change in relative orientation. The other transitions become progressively weaker as the change in relative orientations becomes weaker. This is in complete agreement with the selection principle found for cæsium, except that in bismuth transitions are observed as weak transitions, which are not observed at all in cæsium. But these weak transitions in cæsium may well be present although they are experimentally unobserved; for they are very much too close to the strong transitions (0·01 cm.$^{-1}$) to be resolved from them. The selection principle is therefore found to be the same in the two spectra. The clear physical meaning of the principle (the greater the change in the relative orientations of the nucleus and electron during the transition, the less probable the transition) is most satisfactory.

It remains to consider quantitatively the change of the energy levels of the terms from the nuclear spin. The electron, considered as a point charge, rotating round the nucleus, produces a magnetic field. The nucleus possesses a magnetic moment by virtue of being a rotating charge · it will therefore suffer an increase, or decrease (according to the sense of the rotation), in potential energy, just as the electron does with its magnetic moment in similar circumstances.

The rotation of the electron produces a field H the lines of force being perpendicular to the plane of the orbit; the value of this field a in the locality of the nucleus is

$$\mathrm{H} = \frac{e}{c} \cdot \frac{[r \, . \, v]}{r^3},$$

where r = distance of electron from nucleus, v = velocity of electron. Now the electronic motion is quantised, so that the value of its angular momentum is given by the equation

$$hk/2\pi = m \, [v \, . \, r],$$

so that $\mathrm{H} = ehk/mc \, . \, 2\pi r^3$.

The effect of this magnetic field is to give the nucleus an average potential energy W_k.

$$\mathrm{W}_k = \bar{\mathrm{H}} \, . \, \mathrm{M}_k,$$

M_k being the magnetic moment of the nucleus, $\bar{\mathrm{H}}$ the time average value of the field.

The sign is positive or negative, according to the relative directions of the nuclear spin and the electronic orbital motion. The change in potential energy is thus

$$\Delta W_k = \frac{ehk}{m\,.\,c\,.\,2\pi}\overline{\left(\frac{1}{r^3}\right)}M_k,$$

and the corresponding change in frequency

$$\Delta N_k = \frac{ek}{mc2\pi}\overline{\left(\frac{1}{r^3}\right)}M_k.$$

In order to obtain a numerical solution to this equation, it would be necessary to have an exact knowledge of the electronic orbit so that the function $(\overline{1/r^3})$ could be evaluated. Anything beyond a very rough approximation seems impossible with our present knowledge.

For the *p* orbits an indirect method of calculating this function is available, owing to the fact that it is involved in the equation for the electron spin doublet, so that, by making use of the observed frequency differences due to the alternative directions of the spin in the *p* orbits, this quantity can be eliminated. Unfortunately, this method is confined to the smaller differences of levels which occur in the *p* levels. It cannot be extended to the larger difference, in the *s* level, since the electron spin in the *s* orbit is always in the same direction as the orbital motion, and there is no spin doublet in this orbit.

The electron spin causes the electron to possess a magnetic moment; by virtue of this magnetic moment the electron is acted on by the magnetic field created by its orbital rotation. In order to simplify the calculation of this field the electron is considered as a fixed, but spinning point, charge, and the nucleus as charge rotating round it, in the same orbit and with the same velocity as that usually ascribed to the electron. In this way an equation is found for the change in energy of the electron in a manner very similar to that for the spinning nucleus.

The corresponding change in frequency is thus $\Delta N_e = Zek\,(\overline{1/r^3})\,M_e/mc2\pi$; where M_e is the magnetic moment of the electron according to Goudsmid and Uhlenbeck; but L. H. Thomas* has pointed out that the value of the separation found by this method is (owing to the inexact treatment of the kinematic problem in ascribing to the nucleus the velocity and orbit of the electron) in error by a factor amounting to approximately 0·5, so that the final value for the alkali doublet separation is given by $\Delta N_e = \frac{1}{2}Zek\,(\overline{1/r^3})\,M_e/2\pi mc$.

We now have two equations, the one giving the difference in the level of atom caused by the nuclear spin, and the other giving the difference caused by the electronic spin. The first of these should correspond to the difference in the p levels in the two types of cæsium atom; that is half the difference in the separations of the α and β $1s - m\,{}^2p_{3/2}$ lines and the α and β $1s - m\,{}^2p_{1/2}$ lines, for in the first of these the p separation is subtracted from the s separation (see diagram of energy levels), while in the second it is added to the s separation (in accordance with the fact that the rotation of the electron in the s and ${}^2p_{3/2}$ orbits is in the same direction, while in the ${}^2p_{1/2}$ orbits it is in the opposite direction to these).

---

* 'Nature,' vol. 117, p. 514 (1926).

The second of these two equations gives the difference in the term levels due to the direction of the electron spin, and should correspond to the separation of the terms, *i.e.*, the ordinary alkali doublet separation.

Before numerical values are introduced, attention may be drawn to an approximation which was made in the treatment of the effect of the nuclear spin. Here the field due to the electron was calculated on the assumption that it was a rotating point charge, its spin being neglected; this introduces small errors in the calculated value of the magnetic field created by the electron; the ratio of the true value to the value calculated is of the same order as the ratio of the ${}^2p_{3/2}$ level to the ${}^2p_{1/2}$ level. This error is therefore quite unimportant, being of the order of 1 per cent.

We therefore obtain the following ratio for the nuclear spin separation to the electron spin separation

$$\frac{\Delta N_k}{\Delta N_e} = \frac{ek\,\overline{(1/r^3)}/2\pi mc}{\frac{1}{2}Zek\,\overline{(1/r^3)}/2\pi mc} \cdot \frac{M_k}{M_e} = \frac{1}{\frac{1}{2}Z} \cdot \frac{M_k}{M_e}.$$

If the relative distribution of the charge and mass were the same in the electron and the nucleus, the ratio of their magnetic moments to their mechanical moments would be the same as the ratio of their charges to their masses. Since the nucleus has been assumed to have half quantum and the electron half quantum, the following ratio for the separation of the doublets would be found if the nucleus and electron had the same charge and mass distributions.

$$\frac{\Delta N_k}{\Delta N_e} = \frac{1}{\frac{1}{2}Z} \cdot \frac{Ze/132}{e/1/1840} = \frac{1}{120{,}000}.$$

The experimental value for this ratio is

$$\Delta N_k/\Delta N_e = (0{\cdot}01 \pm 0{\cdot}002)/550 = 1/55000 \pm 20 \text{ per cent.}$$

For the $1s - 3p$ and $1s - 4p$ doublets it is rather greater; but the probable error here is very much larger, and well outside the discrepancy. The value is therefore taken from the $1s - 2p$ lines and found to be approximately 1/60,000, *i.e.*, about twice as great as the value calculated

The experimentally found separation of the p levels is thus about twice as great as the above theory leads one to expect. There are two explanations to this. The positive charge on the nucleus may be situated nearer the periphery, the centre being more nearly neutral; this would increase the magnetic moment; for the velocity of rotation is determined by the moment of inertia of the nucleus, and if the charge is situated farther from the centre than the mass there would follow an increase in the magnetic moment, for a fixed moment of one half quantum. Rutherford from radioactive considerations considers that the charge should be distributed in this way.

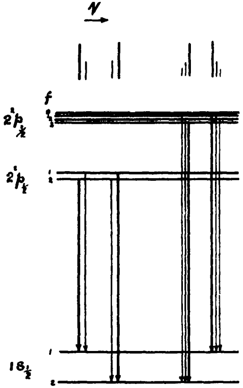


Fig. 4.

There is, however, another possibility; the nucleus may have a spin of $1\frac{1}{2}$ quanta. This would cause the p level to be quadruple; the energy levels and transitions are shown in fig. 4. Unfortunately, the weak components would be too close to be resolved from the strong components, and the separation of these p levels is too small to be observed directly in the lines of the diffuse series; so that the question cannot be solved by directly examining the lines. The only possibility is to observe the Zeeman effect of the principal series doublets. If a field of such strength as to give a Zeeman effect of about twice the separation of the *s* level were applied, each of the usual anomalous Zeeman effect lines should be split into $2n+1$ components (compare Back and Goudsmid). For quantum number $\frac{1}{2}$, this would be two and for $1\frac{1}{2}$, four.

The author takes this opportunity for thanking Prof. Lindemann for the encouragement and advice which he gave throughout the course of this research.

The Internal Conversion of Gamma-Rays.—Part II.

By Bertha Swirles, M.A., Hertha Ayrton Research Fellow, Girton College, Cambridge.

(Communicated by R. H Fowler, F R S – Received August 8, 1928.)

§ 1. *Introduction.*

In a previous paper* the absorption of γ-rays in the K–X-ray levels of the atom in which they are emitted was calculated according to the Quantum Mechanics, supposing the γ-rays to be emitted from a doublet of moment $f(t)$ at the centre of the atom. The non-relativity wave equation† derived from the relativity wave equation for an electron of charge $-\epsilon$ moving in an electro-magnetic field of vector potential **K** and scalar potential V is

$$h^2\nabla^2\phi + 2\mu\left(ih\frac{\partial}{\partial t} + \epsilon \mathrm{V} + \frac{ih\epsilon}{\mu c}(\mathbf{K}\,.\,\mathrm{grad})\right)\phi = 0. \qquad (1)$$

Suppose, however, that **K** involves the space co-ordinates. Then,

$$(\mathbf{K}\,.\,\mathrm{grad})\,\phi \neq (\mathrm{grad}\,.\,\mathbf{K})\,\phi,$$

and the expression $(\mathbf{K}\,.\,\mathrm{grad})\,\phi$ is not Hermitic. Equation (1) cannot therefore be the correct non-relativity wave equation for a single electron in an electro-magnetic field, and we must substitute

$$h^2\nabla^2\phi + 2\mu\left(ih\frac{\partial}{\partial t} + \epsilon \mathrm{V}\right)\phi + \frac{ih\epsilon}{c}\left((\mathbf{K}\,.\,\mathrm{grad})\,\phi + (\mathrm{grad}\,.\,\mathbf{K})\,\phi\right) = 0. \qquad (2)$$

---

* 'Roy. Soc. Proc.,' A, vol. 116, p. 491 (1927). Quoted as *loc. cit.*

† O. Klein, 'Z. Physik,' vol. 41, p. 407 (1927)

Equation (5) (*loc. cit.*) was derived from (1), and therefore needs a further term

$$-\tfrac{1}{2}\cdot\frac{ih\varepsilon}{c^2}\frac{\partial}{\partial z}\left(\frac{1}{r}f'(t-r/c)\right) \tag{3}$$

on the right-hand side.

The expression (13) (*loc. cit.*) therefore becomes

$$\frac{3c^3}{4h\nu^3}\left|\begin{array}{l}\displaystyle\int\left\{\left(\varepsilon-\frac{h\nu\varepsilon}{2c^2\mu}\right)\frac{\partial}{\partial z}\left(\frac{1}{r}\right)e^{i\nu r/c}\,\psi_s\psi_r\right.\\[2ex] \displaystyle\qquad+i\nu\left(\varepsilon-\frac{h\nu\varepsilon}{2c^2\mu}\right)\cdot\frac{1}{cr}\cdot\frac{\partial r}{\partial z}\cdot e^{i\nu r/c}\,\psi_s\psi_r\\[2ex] \displaystyle\qquad\left.-\frac{\nu h\varepsilon}{\mu c^2}\cdot\frac{1}{r}\cdot\frac{\partial\psi_s}{\partial z}\,\psi_r e^{i\nu r/c}\right\}dv\end{array}\right|^2\times M\,d\nu$$

$$=\frac{3c^3\varepsilon^2}{4h\nu^3}\cdot\alpha_{\nu z}\,.\,M\,d\nu,\quad\text{say.} \tag{4}$$

Using this corrected expression to find the absorption coefficient α_ν, where, as before,

$$\alpha_\nu = c^3\varepsilon^2(\alpha_{\nu x}+\alpha_{\nu y}+\alpha_{\nu z})/4h\nu^3, \tag{5}$$

we find the following results instead of those given in the previous paper for the absorption in the K-levels.

Approximation for Wave-Lengths near the Limit.

$$\alpha_\nu \simeq 0\cdot 2\,(\lambda/\lambda_\kappa)^{2\cdot 40}. \tag{6}$$

Approximation for very Short Wave-Lengths.

$$\alpha_\nu \simeq 3\cdot 37\,(\lambda/\lambda_\kappa)^{3\cdot 5}\left[=\frac{4\kappa^2}{2\pi\sigma^2}\cdot 1\cdot 345\,(\lambda/\lambda_\kappa)^{3\cdot 5}\right]. \tag{7}$$

The values of $\alpha_{\nu\kappa}$ for three values of λ/λ_κ in the range investigated experimentally by Ellis and Wooster are given in the following table:—

Table I.

λ/λ_κ.	Ellis and Wooster $\alpha_{\nu\kappa} = 4\cdot 104\,(\lambda/\lambda_\kappa)^{2\cdot 85}$	$\alpha_{\nu\kappa}$ from (5).
0·10	0 0092	0·0006
0·20	0·0576	0·0019
0·30	0·1688	0·0060

§ 2. *The Absorption in the L-Levels.*

In principle the absorption in the L-levels can be calculated in the same way as that in the K-levels. Technically, however, it involves considerably more difficulty and very complicated expressions are obtained. Experimentally, Ellis and Wooster have found that the absorption in the L_I levels is much greater than that in the L_{II} and L_{III} levels. The calculations for the L_I absorption will now be given, and an approximation for any k when $k = n - 1$ for very short wave-lengths. The general calculations for the L_{II} and L_{III} absorption are extremely laborious, and the approximation given is sufficient to indicate that the absorption in the L_{II} and L_{III} levels is much less than that in the L_I levels in agreement with experiment.

Absorption in the L_1 Levels.

Here

$$\psi_s \equiv \psi(2, 0, 0) = 2\gamma^{3/2} e^{-\gamma r}(1 - \gamma r)/2\sqrt{\pi}, \tag{8}$$

ψ_s is already normalised, and

$$\frac{\partial\psi(2, 0, 0)}{\partial x} = 2\gamma^{5/2} e^{-\gamma r}(\gamma r - 2)\frac{\partial r}{\partial x}, \tag{9}$$

with similar expressions in y and z. The (θ, ϕ) integrals in α_{rx}, α_{ry}, α_{rz} are the same as in the K absorption, and we have again $k' = 1$.

Then, multiplying by 2, since there are $2L_I$ electrons in the atom,

$$\alpha_\nu = \frac{c^3\epsilon^2}{4h\nu^3} \cdot 8\gamma^3\xi'^2 \left| \begin{array}{l} \displaystyle\int_0^\infty \Big\{-\left(1 - \frac{b\kappa}{2}\right) r(1 - \gamma r) \\ \displaystyle\qquad + i\beta\left(1 - \frac{b\kappa}{2}\right) r^2 (1 - \gamma r) \\ \displaystyle\qquad - b\kappa\gamma r^2 (\gamma r - 2)\Big\} \\ \displaystyle\qquad \times e^{(i\beta - \gamma)r} \int_C e^{zr} (z - \gamma')^{1 - in'} (z + \gamma')^{1 + in'}\, dz\, dr \end{array} \right|^2, \tag{10}$$

where the notation is the same as in *loc. cit.* and

$$\gamma = \alpha/2.$$

Therefore (10) becomes

$$\frac{c^3\epsilon^2\alpha^3\xi'^2}{4h\nu^3} \left| \begin{array}{l} \displaystyle -\left(1 - \frac{b\kappa}{2}\right)A' + \alpha\left(\tfrac{1}{2}\left(1 + \frac{3b\kappa}{2}\right) + ib\left(1 - \frac{b\kappa}{2}\right)\right)B' \\ \displaystyle\qquad - \alpha^2\left(\tfrac{1}{4}b\kappa + \frac{ib}{2}\left(1 - \frac{b\kappa}{2}\right)\right)C' \end{array} \right|^2, \tag{11}$$

where,

$$A' = \int_0^\infty \int_C e^{(i\beta - \gamma + z)r} \cdot r(z - \gamma')^{1 - in'} (z + \gamma')^{1 + in'}\, dz\, dr,$$

$$B' = \int_0^\infty \int_C e^{(i\beta - \gamma + z)r} \cdot r^2 (z - \gamma')^{1 - in'} (z + \gamma')^{1 + in'}\, dz\, dr,$$

$$C' = \int_0^\infty \int_C e^{(i\beta - \gamma + z)r} \cdot r^3 (z - \gamma')^{1 - in'} (z + \gamma')^{1 + in'}\, dz\, dr.$$

Integrating first with respect to r, we obtain integrals, each of which is equal to

$$-2\pi i\,(1-e^{-2\pi n'})\text{ (residue at }(\gamma-i\beta)+\text{residue at }\infty).$$

Integral A′. The residue at $(\gamma-i\beta)$ is

$$\alpha\left(\frac{in'-2+2bn'}{in'+2+2bn'}\right)^{in}(3-2ib). \tag{12}$$

The residue at ∞ is

$$\alpha\,(2ib+1). \tag{13}$$

Integral B′. The residue at $(\gamma-i\beta)$ is

$$-2\left(\frac{in'-2+2bn'}{in'+2+2bn'}\right)^{in'}(1+n'^2)+\left(\frac{in'-2+2bn'}{in'+2+2bn'}\right)^{in'+1}(n'^2+in')$$
$$+\left(\frac{in'-2+2bn'}{in'+2+2bn'}\right)^{in'-1}(n'^2-in'). \tag{14}$$

The residue at ∞ is

$$2. \tag{15}$$

Integral C′. The residue at $(\gamma-i\beta)$ is

$$\frac{128}{\alpha}\,\frac{n'^2\,(1+n'^2)}{(in'-2+2bn')^2\,(in'+2+2bn')^2}\left(\frac{in'-2+2bn'}{in'+2+2bn'}\right)^{in'}. \tag{16}$$

The residue at ∞ vanishes.

For brevity, let

$$\left(\frac{in'-2+2bn'}{in'+2+2bn'}\right)=p. \tag{17}$$

Then (11) becomes

$$\left|\frac{\pi^2c^3\varepsilon^2\alpha^5\xi'^2}{h\nu^3}\,(1-e^{-2\pi n'})^2\right|^2\times$$
$$\times\left|\begin{matrix}p^{in'}\left(-\left(1-\frac{b\kappa}{2}\right)(3-2ib)-2\,(1+n'^2)\right)\left(\tfrac{1}{2}\left(1+\frac{3b\kappa}{2}\right)+ib\left(1-\frac{b\kappa}{2}\right)\right)\\ -\dfrac{128n'^2\,(1+n'^2)}{in'-2+2bn')^2\,(in'+2+2bn')^2}\left(\tfrac{1}{4}b\kappa+\frac{ib}{2}\left(1-\frac{b\kappa}{2}\right)\right)\\ +(p^{in'+1}(n'^2+in')+p^{in'-1}(n'^2-in'))\left(\tfrac{1}{2}\left(1+\frac{3b\kappa}{2}\right)+ib\left(1-\frac{b\kappa}{2}\right)\right)\\ -\left(1-\frac{b\kappa}{2}\right)(2ib+1)+1+\frac{3b\kappa}{2}+2ib\left(1-\frac{b\kappa}{2}\right)\end{matrix}\right| \tag{18}$$

$$=\frac{\mu\varepsilon^2\alpha^2c^3}{64\pi^2h^2\mathrm{R}^3\mathrm{Z}^6}\left(\frac{\lambda}{\lambda_\kappa}\right)^3\frac{n'^2}{n'^2+1}\,(1-e^{-2\pi n'})^{-1}\,|\mathrm{P}|^2,\text{ say.} \tag{19}$$

Only the approximation for very short wave-lengths will be given, as the wave-lengths considered experimentally are short compared with the limiting wave-length for L absorption.

Let

$$p^{in'} = 1 + \alpha n'^2 + \beta n'^4 + \dots \tag{20}$$

$$p^{in'+1} = 1 + \mathrm{A}n' + \mathrm{B}n'^2 + \mathrm{C}n'^3 + \dots \tag{21}$$

$$p^{in'-1} = 1 - \mathrm{A}n' + \mathrm{B}n'^2 - \mathrm{C}n'^3 - 1 \dots \tag{22}$$

where, by Maclaurin's expansion,

$$\alpha = -2i/\sigma, \quad \mathrm{A} = -2/\sigma, \quad \mathrm{B} = -2i/\sigma + 2/\sigma^2, \quad \mathrm{C} = 1/2\sigma + 5i/\sigma^2 - 2/\sigma^3.$$

Now,

$$b = \tfrac{1}{4}\sigma\,(1 + 4n'^{-2}). \tag{23}$$

Therefore,

$$\begin{aligned} \mathrm{P} \simeq {} & (1 + \alpha n'^2 + \beta n'^4 + \dots)\left(-4 - n'^2\left(1 + \frac{3b\kappa}{2}\right) - 2in'^2 b\left(1 - \frac{b\kappa}{2}\right) + 0(n'^2)\right) \\ & + (n'^2(2 + 2i\mathrm{A}) + n'^4(2\mathrm{B} + 2i\mathrm{C}) + \dots)\left(\tfrac{1}{2}\left(1 + \frac{3b\kappa}{2}\right) + ib\left(1 - \frac{b\kappa}{2}\right)\right) \\ & + 2b\kappa. \end{aligned} \tag{24}$$

The first term, which does not vanish, is the constant term, which is equal to

$$-\,2\kappa/\sigma.$$

Therefore,

$$\alpha_\nu \simeq \frac{4\kappa^2}{2\pi\sigma^2}\cdot\frac{1\cdot345}{8}\left(\frac{\lambda}{\lambda_\kappa}\right)^{3\cdot5}, \tag{25}$$

i.e., α_ν again varies as $\lambda^{3\cdot5}$ and the coefficient is exactly one-eighth of that for the K absorption.

The expansion of $p^{in'}$ is valid for values of n' up to 0·25, *i.e.*,

$$\lambda/\lambda_\kappa < 0\cdot06.$$

Absorption in the $\mathrm{L_{II}}$ *and* $\mathrm{L_{III}}$ *Levels.*

For the $\mathrm{L_{II}}$ and $\mathrm{L_{III}}$ levels of a hydrogen-like atom

$$n = 2, \quad k = 1.$$

For the slightly more general atom model considered by Oppenheimer,* n and k are not necessarily themselves integral, although their difference is an integer.

We shall now consider the case when n and k are not necessarily integers and

$$n - k = 1.$$

---

* J. R. Oppenheimer, 'Z. Physik,' vol. 41, p. 268 (1927).

This includes the case of the K-absorption as well as that of the L_{II} and L_{III} absorption. We shall find that it is possible to obtain approximations to the integrals involved for very short wave-lengths.

Now,

$$u_{k+1,\,k,\,m}(r) = (2\gamma r)^k e^{-\gamma r}\,\Gamma(2k+2). \tag{26}$$

This expression differs from (17) (*loc. cit.*) by a constant factor only. Both expressions require normalising.

We obtain integrals of the form

$$\int_0^\infty \int_C r^{k+k'} (z-\gamma')^{k'-in'} (z+\gamma')^{k'+in'} e^{r(z+i\beta-\gamma)}\,dz\,dr$$
$$= \Gamma(k+k'+1) \int_C \frac{(z-\gamma')^{k'-in'} (z+\gamma')^{k+in'}}{(-z-i\beta+\gamma)^{k+k'+1}}\,dz \tag{27}$$

and

$$\int_0^\infty \int_C r^{k+k'+1} (z-\gamma')^{k'-in'} (z+\gamma')^{k+in'} e^{r(z+i\beta-\gamma)}\,dz\,dr$$
$$= \mathrm{T}(k+k'+2) \int_C \frac{(z-\gamma')^{k'+in'} (z+\gamma')^{k'+in'}}{(-z-i\beta+\gamma)^{k+k'+2}}\,dz. \tag{28}$$

Now the integral round the double circuit is equal to

$$(1-e^{2\pi i(k'-in')})\,(1-e^{2\pi i(k'+in)})$$

times a simple integral joining the branch points γ', $-\gamma'$.

$$\int_{-\gamma'}^{\gamma'} \frac{(z-\gamma')^{k'-in'} (z+\gamma')^{k'+in'}}{(-z-i\beta+\gamma)^{k+k'+1}}\,dz \tag{29}$$

$$= n'^{2(k+k'+1)} \int_{-\gamma'}^{\gamma'} \frac{(z-\gamma')^{k'-in'} (z+\gamma')^{k'+in'}}{(-(z-\gamma+i\sigma\alpha/n'^2)n'^2-i\sigma\alpha)^{k+k'+1}}. \tag{30}$$

To obtain an approximation for small n', *i.e.*, large $|\gamma'|$, we take the line integral round a semicircle of radius $|\alpha'|$.

Then (30)

$$\simeq n'^{2(k+k'+1)}\gamma'^{2k'+1} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{(e^{2i\theta}-1)^{k'}\,.\,i\,.\,e^{i\theta}\,d\theta}{(-i\sigma\alpha)^{k+k'+1}}, \tag{31}$$

$$\simeq \frac{n'^{2k+1} i^{k-k+1} \alpha^{k'-k} e^{i\pi k'}}{\sigma^{k+k'+1}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{i\theta}(1-e^{2i\theta})^{k'}\,d\theta. \tag{32}$$

$$\simeq e^{\frac{i\pi}{2}(5k'+2k+4)}\,\frac{n'^{2k+1}\alpha^{k'-k}}{\sigma^{k'+k+1}}\,.\,2\,.\,\sum_\tau \frac{(-1)^\tau}{1+2\tau}\binom{k'}{\tau} = \mathrm{X},\ \text{say}. \tag{33}$$

Similarly,

$$\int_{-\gamma}^{\gamma}\frac{(z-\gamma')^{k-in}\,(z+\gamma')^{k+in}}{(-z-i\beta+\gamma)^{k+k+1}}\,dz$$

$$\simeq e^{\frac{i\pi}{2}(\omega k+2k+\nu)}\,\frac{n^{2k+3}\alpha^{k-k-1}}{\sigma^{k+k+2}}\;\;2\sum_{\tau}\frac{(-1)^{\tau}}{1+2\tau}\binom{k'}{\tau}=\mathrm{Y},\text{ say} \qquad (34)$$

Then,

$$\mathrm{Y}=i\mathrm{X}\;\;n'^2/\alpha\sigma \qquad (35)$$

It is clear that the expressions (33) and (34) do not take singular values when k and k' are integers and

$$k'-k=\pm 1$$

We now consider the special cases

(i) $$n=2,\quad k=1,\quad m=0,$$
$$k'=0,\quad m=0.$$

Then,

$$\psi_s=\psi\,(2,1,0)=\left(\frac{\alpha^5}{4!}\right)^{\frac12}\;re^{-\gamma r}\cos\theta\;\;\tfrac12\sqrt{\frac{3}{\pi}},\text{ already normalised}$$

$$\psi_r=\psi\,(n',0,0)=\int_{-\gamma}^{\gamma}e^{zr}\,(z-\gamma')^{-in}\,(z+\gamma')^{in}\,dz$$

Now from (4) it may be shown by a simple differentiation by parts that

$$\alpha_{\nu s}=\frac{3c^3\varepsilon^2}{4h\nu^3}\left|\;\int\left\{\left(1+\frac{b\kappa}{2}\right)\frac{\partial}{\partial z}\left(\frac{1}{r}\right)e^{i\nu r/c}\psi_s\psi_r\right.\right.$$
$$\left.\left.+\,i\beta\left(1+\frac{b\kappa}{2}\right)\;\frac{1}{r}\,\frac{\partial r}{\partial z}\,e^{i\nu r/c}\psi_s\psi_r+\frac{b\kappa}{2}\;\frac{1}{r}\,\frac{\partial\psi_r}{\partial z}\;\psi_s e^{i\nu r/c}\right\}dv\;\right|^2 \qquad (36)$$

Now

$$\frac{\partial\psi\,(n',0,0)}{\partial x}=\frac{\partial\psi\,(n',0,0)}{\partial r}\,\frac{\partial r}{\partial x}\qquad(x=x,\,y\;\,z)$$

and

$$\frac{\partial\psi\,(n',0,0)}{\partial r}=\int_{-\gamma}^{\gamma}ze^{zr}\,(z-\gamma')^{-in}\,(z+\gamma)^{in}\;dz$$

The sum of the squares of the $(\theta,\,\phi)$ integrals is $1/3$, and again multiplying by 2, since there are two electrons with $n=2$, $k=1$, $m=0$, we have

$$\alpha_\nu=\frac{c^3\varepsilon^2\xi'^2\alpha^5}{4\;\;4!\,h\nu^3}\left|-\left(1-\frac{b\kappa}{2}\right)\mathrm{D}+i\beta\left(1+\frac{b\kappa}{2}\right)\mathrm{E}+b\kappa\mathrm{F}\right|^2, \qquad (37)$$

where,

$$D = \int_{-\gamma'}^{\gamma'} \frac{(z-\gamma')^{-in'}(z+\gamma')^{in'}}{(-z-i\beta+\gamma)^2}\,dz,$$

$$E = 2\int_{-\gamma'}^{\gamma'} \frac{(z-\gamma')^{-in'}(z+\gamma')^{in'}}{(-z-i\beta+\gamma)^3}\,dz,$$

$$F = 2\int_{-\gamma'}^{\gamma'} \frac{z(z-\gamma')^{-in'}(z+\gamma')^{in'}}{(-z-i\beta+\gamma)^3}\,dz = (\gamma - i\beta)\,E - 2D.$$

Therefore (37)

$$= \frac{c^3\epsilon^2\xi_\gamma'^2\alpha^5}{4\,.\,4\,!\,h\nu^3}\left| -D\left(1+\frac{5b\kappa}{2}\right) + E\alpha\left(\tfrac{1}{2}b\kappa + ib\left(1-\frac{b\kappa}{2}\right)\right)\right|^2 \tag{38}$$

$$b = \frac{\sigma}{4}(1 + 4n'^{-2})$$

Now,

$$\xi'^2 = \frac{\dfrac{2\mu}{h}\,|2\gamma'|^{-1}\,e^{3\pi n'}}{\Gamma(1-in')\,\Gamma(1+in')}$$

$$= \frac{\mu e^{4\pi n'}}{2\pi h\alpha}\left(1 - e^{-2\pi n'}\right). \tag{39}$$

Now from (33) and (34)

$$D \simeq e^{3i\pi}\,.\,2n'/\alpha\sigma^3, \tag{40}$$

$$E \simeq ie^{3i\pi}\,.\,4n'^2/\alpha^2\sigma^3. \tag{41}$$

Therefore (38),

$$\simeq \frac{\mu\epsilon^2\alpha^2c^3}{8\pi^3\,.\,4\,.\,4\,!\,h^2R^3Z^6}\cdot\frac{9\kappa^2}{\sigma^2}\left(\frac{\lambda}{\lambda_\kappa}\right)^{4\cdot 5}.$$

Therefore,

$$\alpha_\nu \simeq 0\cdot 17\,(\lambda/\lambda_\kappa)^{4\cdot 5}, \tag{42}$$

(ii) $n = 2,\ k = 1,\ m = 0, \pm 1,\ \psi(2, 1, m) = re^{-\gamma r}\cos\theta e^{im\phi}.$

Then m' must equal m and

$$\alpha_{\nu z} = \frac{3c^3\epsilon^2\xi_{\gamma r}{}^2\xi_{\gamma r}{}'^2\xi_{\theta}{}^2\xi_{\theta}{}'^2}{4h\nu^2}\times$$

$$\left|\int\left\{-\left(1-\frac{b\kappa}{2}\right).\frac{1}{r^2}.re^{-\gamma r}P_1{}^{|m|}(\cos\theta)\cos\theta + i\beta\left(1-\frac{b\kappa}{2}\right).\frac{1}{r}.re^{-\gamma r}P_1{}^{|m|}(\cos\theta)\cos\theta \right.\right.$$
$$- b\kappa.\frac{1}{r}\left(\left(e^{-\gamma r}-\gamma re^{-\gamma r}\right)P_1{}^{|m|}(\cos\theta)\cos\theta - \frac{1}{r}.re^{-\gamma r}\left(|m|\cos\theta P_1{}^{|m|}(\cos\theta) - P_1{}^{|m|+1}(\cos\theta)\right)\right)\Bigg\}$$
$$\left.\times\, e^{i\phi r}P_{k'}^{|m|}(\cos\theta)\sin\theta\, u_{\kappa'}(r)\, r^2\, dr\, d\theta.\right|^2 \tag{43}$$

Now

$$\int_0^\pi (\mathrm{P}_1{}^{|m|}(\cos\theta)\cos\theta - |m|\mathrm{P}_1{}^{|m|}(\cos\theta)\cos\theta + \mathrm{P}_1{}^{|m|+1}(\cos\theta)\sin\theta) \\ \times \mathrm{P}_{k'}^{|m'|}(\cos\theta)\sin\theta\, d\theta$$

vanishes if $|m| = 0$ or 1, and the other integrals vanish unless

$$|k' - k| = 1.$$

We have already considered $k = 0$. We now take $k = 2$. By a similar consideration of $\alpha_{\nu x}$, $\alpha_{\nu y}$, we find that

$$\alpha_\nu = \frac{c^3\varepsilon^2}{6h\nu^3}\xi_{sr}{}^2\xi_{sr}{}'^2 \left| \int_0^\infty \int_{-\gamma'}^{\gamma'} \left(-\left(1 - \frac{b\kappa}{2}\right)r^3 + \left(i\beta\left(1 - \frac{b\kappa}{2}\right) + b\kappa\gamma\right)r^4\right) \times \right. \\ \left. \times\, e^{(i\beta-\gamma+s)r}(z-\gamma')^{2-in'}(z+\gamma')^{2+in'}\,dr\,dz \right|^2 \tag{44}$$

$$= \frac{c^3\varepsilon^2}{6h\nu^3}\xi_{sr}{}^2\xi_{sr}{}'^2 \left| -\mathrm{D}'\left(1 - \frac{b\kappa}{2}\right) + \mathrm{E}'\left(i\beta\left(1 - \frac{b\kappa}{2}\right) + \frac{b\kappa\alpha}{2}\right) \right|^2 \tag{45}$$

where

$$\xi_{sr}{}^2 = \frac{\alpha^5}{4\,!}, \qquad \xi_{sr}{}'^2 = \frac{\mu n'^4}{2^7\pi h\alpha^5}e^{4\pi n'}(1 - c^{-2\pi n'}).$$

$$\mathrm{D}' = 3\,!\int_{-\gamma'}^{\gamma'} \frac{(z-\gamma')^{2-in'}\,(z+\gamma')^{2+in'}}{(-z-i\beta+\gamma)^4}\,dz,$$

$$\mathrm{E}' = 4\,!\int_{-\gamma'}^{\gamma'} \frac{(z-\gamma')^{2-in'}\,(z+\gamma')^{2+in'}}{(-z+i\beta+\gamma)^5}\,dz.$$

From (33) and (34)

$$\mathrm{D}' \simeq 3\,!\,n^3\alpha/\sigma^4, \quad \mathrm{E}' \simeq i\,.\,4\,!\,n'^5/\sigma^5,$$

and after some reduction we find

$$\alpha_\nu \simeq 5{\cdot}50\,(\lambda/\lambda_\kappa)^{6\cdot5}. \tag{46}$$

§ 3. *Discussion of Results.*

The main features of the experimental results obtained by Ellis and Wooster for γ-rays of wave-lengths in a range

$$0{\cdot}03 < \lambda/\lambda_\kappa < 0{\cdot}5 \tag{47}$$

are that $\alpha_{\nu\kappa}$ varies with λ to a power of about 3, and is of the order given in Table I; further, that the ratio of the $\mathrm{L_I}$ to the K absorption is about 1/5, and the absorption in the $\mathrm{L_I}$ levels is greater than that in the $\mathrm{L_{II}}$ and $\mathrm{L_{III}}$ levels.

From (6) and (7) we see that, theoretically, the K absorption coefficient varies with λ to a power not widely differing from 3 in the limiting cases considered, and from Table I that $\alpha_{\nu\kappa}$ varies with λ to a power between 2 and

3, but is of the order of 1/10 too small. From (25) and (7) the absorption in the L_I levels is 1/8 of that in the K-levels, and the L_{II} and L_{III} absorption is from (42) and (46) much smaller than the L_I absorption for small wave-lengths.

To sum up, we may say that with regard to the variation of α_ν with λ, and the ratio of the absorption in the K and L_I levels and in the different L levels, the present theory is reasonably in agreement with experiment. Further, by comparison with Oppenheimer's results for the photo-electric effect, the theory indicates that the ratio between internal and external absorption varies only slightly over a wide range of wave-lengths in agreement with experiment. On the other hand, the order of magnitude of the theoretical internal absorption coefficient is everywhere of the order of 1/10 too small.

It is hoped to carry out the calculations with a better approximation to the characteristic functions of a heavy atom than that given by those of a hydrogen atom.

The Increase of Thermionic Currents from Tungsten in Strong Electric Fields.

By Russell S. Bartlett, Ph.D., Yale University.

(Communicated by O. W. Richardson, F.R.S.—Received August 10, 1928.)

Schottky[*] has pointed out that there should be an increase in thermionic currents above the saturation value with the application of strong electric fields. His theoretical expression is based on the assumption that an electron is prevented from escaping by the attraction of its electric image in the emitting wire, and that the external field neutralises part of this image force field. In the same paper he reports an approximate experimental verification of his theoretical law, though the details are lacking. Since that paper, reference has been made to this increase in current, and it has been applied as a correction in determining the thermionic work function. Becker and Mueller[†] have discussed the problem, but apparently assume the correctness of Schottky's expression. Under the circumstances it seemed that an experimental investigation was needed, all the more because Schottky's theoretical expression seems open to question on certain points.

---

[*] 'Phys. Z.,' vol. 15, p. 872 (1914).

[†] 'Phys. Rev.,' vol. 31, p. 431 (1928).

*Theoretical Considerations.**

Let us suppose that a single electron is escaping from a plane electrode, and that the image force alone is acting upon it.

$$\frac{dV}{dx}=\frac{e}{x^2},\qquad V=\frac{e}{4}\int_{x_0}^{x}\frac{dx}{x^2}=\frac{e}{4x_0}-\frac{e}{4x}.$$

Suppose now an external field E is applied, in a direction to assist the escape of the electron.

$$V'=\frac{e}{4x_0}-\frac{e}{4x}-Ex,\qquad \frac{dV'}{dx}=\frac{e}{4x^2}-E.$$

And, for the potential minimum,

$$x_1=\tfrac{1}{2}\sqrt{\frac{e}{E}},\qquad V_1'=\frac{e}{4x_0}-\frac{e}{2\sqrt{e/E}}-\frac{E}{2}\sqrt{\frac{e}{E}}=\phi_0-\sqrt{Ee}.$$

Putting this expression in the thermionic emission equation, we obtain

$$I=AT^2\epsilon^{\frac{-\phi_0}{kT}+\frac{\sqrt{Ee}}{kT}}=I_0\epsilon^{\frac{\sqrt{Ee}}{kT}},\qquad I=I_0\epsilon^{\frac{4\cdot 38\sqrt{E}}{T}}.$$

Where E is expressed in volts per centimetre.

For fields between 300 and 30,000 volts per centimetre, the distance of the potential minimum from the surface ranges from 10^{-5} to 10^{-6} cm. With concentric cylinders for electrodes, the departure from the plane electrode case will be negligible if the radius of the emitting wire is greater than 10^{-3} cm. From the dimensions of the electrodes it is possible to calculate the electric field at the cathode in terms of the applied voltage, if the effect of space charge is neglected. As a second approximation we may take $E=C(V-V_0)$ where V_0 is the voltage necessary to produce saturation. Referring now to Schottky's expression, we find that

$$\log I/I_0=4\cdot 38\sqrt{E/T}=4\cdot 38\sqrt{C(V-V_0)}/T.$$

So plotting $\log I/I_0$ against $\sqrt{V-V_0}$ should give a straight line if the image force is the only force effective.

Experimental Procedure.

The experiments were performed with a tungsten wire, concentric in a molybdenum cylinder, and maintained in position by a small bow spring at each end, the whole mounted in a quartz tube which permitted thorough bathing in an electric furnace. The heating current for the tungsten cathode was furnished by accumulators of sufficient capacity to maintain the temperature sensibly constant throughout a series of observations. The voltage between electrodes was furnished by banks of cells amounting to about 1000 volts in all. The multiplying factor introduced by the dimensions of the apparatus gave maximum fields between 30,000 and 60,000 volts per centimetre for the different wires used. The thermionic currents were measured by galvanometers and microammeters, suitably shunted. Filament temperatures were determined from the resistance, using Worthing and Forsythe's* data, and checked by the saturation thermionic emission currents.

---

* Based on Schottky's derivation, *loc. cit.*

A considerable amount of manipulation was required before the following satisfactory procedure was devised. After preliminary out-gassing in the electric furnace, the tungsten wire was glowed out at about 3000° K. Then the cylindrical anode was raised by bombardment to a bright red. After this it was possible to maintain the vacuum at better than 10^{-6} mm. of mercury. But to obtain consistent and reproducible results it was necessary to repeat the glowing and bombardment before each series of observations. When that had been done, measurements of the thermionic emission current were made at intervals of about 40 volts in the potential between electrodes.

Experimental Results.

Satisfactory results were obtained with two wires: A, ordinary commercial tungsten, not thoriated, of 0·08 mm. diameter, and B, specially pure tungsten, obtained through the kindness of the General Electric Company Research Laboratories, of 0·095 mm. diameter. Runs were made at five different temperatures with A, and at eleven with B, between 1500° and 2400° K. Figs. 1 to 4 show typical results obtained with these wires—first I plotted against V, and then log I against $\sqrt{V}$ and against $\sqrt{V - V_0}$. For all runs the curves obtained were of the same general form.

According to the image force theory, the points ($\log I \sim \sqrt{V - V_0}$) should lie on a straight line. Though the space charge correction, substituting $V - V_0$ for V, brings the points into much better alignment, it was impossible to select any value for V_0, even absurdly large, which would completely eliminate the curvature. A close study of all the observations forces one to the conclusion that Schottky's expression does not quite fit the facts. Certainly the residual curvature is greater than can be accounted for by experimental error.

---

* 'Phys. Rev.,' vol. 18, p. 144 (1921).

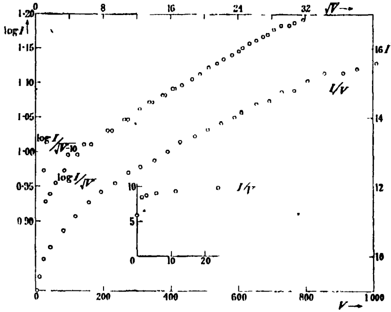


FIG. 1.—Series A 11. T = 1900° K. $I_s = 3 \cdot 5 \times 10^{-6}$ amp./sq. cm.

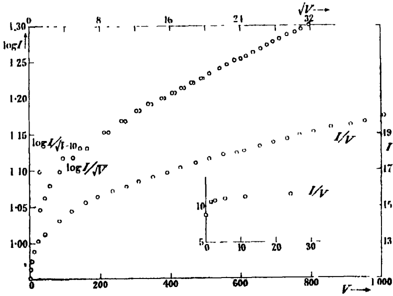


FIG. 2.—Series A 12. T = 1945° K. $I_s = 9 \cdot 2 \times 10^{-6}$ amp./sq. cm.

However, it is still possible to obtain a good value for the average slope of the log $I \sim \sqrt{V - V_0}$ curves, from which we can investigate the variation of this slope with temperature. The theory predicts that it should be inversely proportional to T, and fig. 5 shows to what extent this is borne out by observation. In interpreting these points due weight must be given to the experimental difficulties and uncertainties. In the first place, the heat treatment caused a certain amount of wasting of the tungsten wire between runs,

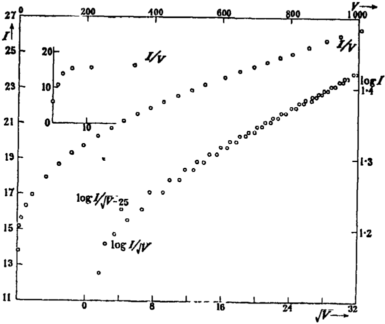


FIG 3.—Series B 6. T = 2040° K. $I_0 = 1\cdot27 \times 10^{-3}$ amp./sq. cm.

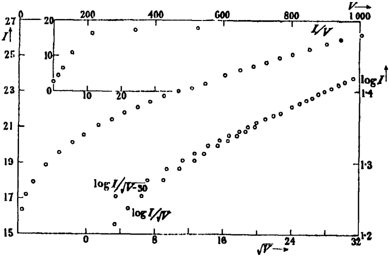


FIG. 4.—Series B 8. T = 2200° K. $I_0 = 9\cdot94 \times 10^{-3}$ amp./sq. cm.

which in turn introduces some uncertainty into the determination of temperature and electric field. Further, at the lowest temperatures it was necessary to correct for an insulation leak inside the tube, while at the highest temperature an error may be introduced by the effect of space charge.

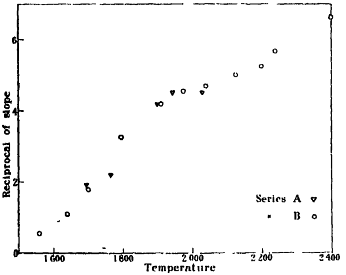


Fig. 5.

It is seen from the figure that the change of slope with temperature is in the right direction, but much greater than that predicted by theory. Even if the end points are discredited to some extent for the reasons given above, it is still impossible to reconcile the results with the theory.

Quantitatively these results agree with the theory at about 1900° K. For higher temperatures the experimental slope is too small by amounts varying up to 25 per cent., while for lower temperatures the departures are much greater, and in the opposite direction, the experimental slope being about six times the theoretical at 1560° K.

In thermionic emission work questions are always asked as to the purity of the tungsten, and the possibility of surface contaminations. It was partly to settle that difficulty that the vigorous heat treatment was used, and as a guarantee of the purity of the tungsten, the temperatures calculated from emission data were consistently lower than those calculated from the resistance, showing that the emission was, if anything, less than what is normal for pure tungsten. The result of that heat treatment is perhaps best illustrated by a comparison of figs. 6 to 7 for runs taken with filament A without any vigorous heating just before the observations, with similar runs after heat treatment

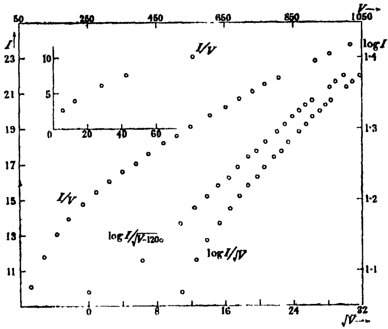


FIG. 6.—A 5. T = 1750°. $I_0 = 3\cdot86 \times 10^{-3}$.
Filament not glowed out.

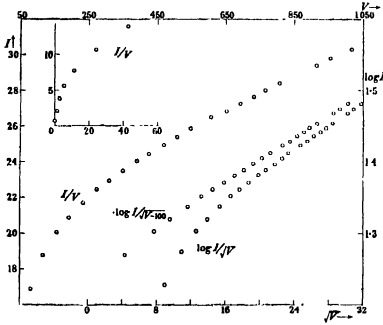


FIG. 7.—A 6. T = 1970°. $I_0 = 6\cdot86 \times 10^{-3}$.
Filament not glowed out.

(figs. 1 and 2). It is noticeable that without the heat treatment there is nothing that can be called a saturation current. Similar, though less marked, results were obtained for filament B, the especially pure tungsten.

Despite this lack of saturation, it is still possible to bring these points into fairly good agreement with theory, by choosing V_0 of the order of 100 volts.

The insulation and gaseous discharge leak inside the tube was measured immediately following each series of observations, and corrected for where necessary. This leak amounted to about 20 per cent. of the thermionic current at the lowest temperature, and was, of course, entirely negligible at higher temperatures.

Discussion of Results.

It was pointed out above that there is no correction necessary for the cylindrical electrodes in this experiment. Further, account is taken, to a good degree of approximation, of the effect of space charges in diminishing the applied field. Even then Schottky's theoretical expression is not exactly verified in observations at constant temperature, still less so, apparently, when it comes to variations of the effect with temperature. A probable reason for this discrepancy is not far to seek.

The theoretical expression was derived on the assumption that a single electron was moving out from the metal surface acted upon by the attraction to its electric image in the metal. Though the electrons escaping from the metal are comparatively few in number, there must be a great stream, of 10^{20} per square centimetre per second or more, passing through the surfaces and being turned back by the work function force. We saw above that for the range of this experiment the work function extends out from the surface a distance of 10^{-5} to 10^{-6} cm. So that for the electron just escaping there are other electrons as close as is the surface charge, and these electrons themselves, and their images in the metal surface, must have some effect upon the particular electron under consideration. An attempt has been made upon this problem, somewhat analogous to the problem of space charge, but so far it has yielded nothing in as close agreement with experiment as the Schottky equation.

Summary.

Experimental results for the increase of thermionic currents with applied electric field at constant temperature show general agreement with theory, but the departures from a predicted straight line are greater than can be accounted for by experimental deficiencies. In the dependence of this rate of increase upon temperature the failure of experiment to agree with the theory is still more marked, even after due allowance is made for certain experimental difficulties.

It is suggested that Schottky's equation should be modified to take account of the influence of neighbouring electrons close to the surface—electrons that emerge from the surface, but do not completely escape.

Attention is drawn to the marked effect of surface impurities in the cathode upon the experimental results.

The author wishes to express his thanks to the laboratory staff at King's College, London, where this work was carried out, for their kindness and assistance; to Dr. A. J. Waterman for many helpful discussions; and in particular to Prof. Richardson for his interest and encouragement.

On the Rate at which Particles take up Random Velocities from Encounters according to the Inverse Square Law.

By L. H. Thomas, Ph.D., Trinity College, Cambridge.

(Communicated by R. H. Fowler, F.R.S.—Received August 17, 1928.)

Summary.

This paper deals theoretically with the changes in the velocities of particles which move through a cloud of particles interacting according to the inverse square law. The formulæ obtained—for low densities—lead to "effective mean free paths" shorter than might have been expected. It seems not impossible that the rapid rate at which beams of electrons moving through highly ionised gases have been observed by Langmuir[*] to take up Maxwellian velocity distributions may be explained in this way.

1. *Introduction.*

This problem has been discussed by Jeans[†] for stellar encounters. He takes into account all the encounters for which the closest distance of approach is less than the normal distance between adjacent stars and neglects the remaining encounters. These latter have, however, in his case, at least as great an effect as the former. Such a correction, while unimportant in cosmogony, is important in atomic problems, where it may be larger, and it is necessary to consider these distant encounters in detail.

---

[*] 'Phys. Rev.,' vol. 26, p. 585 (1925).

[†] 'Astronomy and Cosmogony,' p. 309.

The condition that limits the effectiveness of distant encounters is not that used by Jeans—that if several stars are within a distance a at the same instant, their effects tend to neutralise one another. This neutralisation is already taken care of by our finding the root mean square change of momentum. During an encounter the changes in the velocities of the stars are partly due to their mutual attraction and partly due to their attractions by other stars. The true condition that the encounter may be effective is that the latter part of this change is small, not necessarily compared with the former, but compared with their whole velocity, *i.e.*, that the closest distance of approach is smaller than the mean free path of either star

Although the idea of a mean free path, inaccurate at best, is still less accurate where there are inverse square forces, in this problem it is only required to fix the argument of a logarithm and, for low densities, its inaccuracy will affect the final result little. Instead, therefore, of considering all the particles as having their velocities gradually changed, a fraction of them are supposed, for this purpose, to have, in any interval of time, their velocities completely changed.

It must be observed that the effective mean free path for an encounter so obtained depends on the mean free paths of both encountering particles. It should also be observed that whereas a particle moving among others at rest can only lose speed, one moving among others with one-tenth its energy has a not negligible chance of gaining speed, especially if the latter are much more massive.

The encounters are divided into close encounters, for which the mean free paths are disregarded, and distant encounters which cause separately only small changes of velocity, and in each case the mean transfer of momentum and the mean square transfer of momentum in unit time are calculated.

2. *Close Encounters.*

Suppose there are 8N particles (2) per unit volume of mass m_2 moving with velocity v in the ξ direction relative to a particle (1) of mass m_1 and attracting or repelling it at distance r with force

$$\frac{\lambda}{r^2} = \frac{\mu}{r^2}\frac{m_1 m_2}{m_1 + m_2}. \tag{2.1}$$

If p is the perpendicular distance of the line of relative motion of one of the former from the latter, the angle of deflection in the relative orbit, δ, is given by

$$1 + \frac{p^2 v^4}{\mu^2} = \frac{2}{1 - \cos\delta}. \tag{2.2}$$

If the plane of the relative orbit makes angle ε with the (ζ, η) plane, the momentum transferred in the encounter is

$$\left(\frac{m_1 m_2}{m_1 + m_2} v (1 - \cos\delta), \quad \frac{m_1 m_2}{m_1 + m_2} v \sin\delta \cos\varepsilon, \quad \frac{m_1 m_2}{m_1 + m_2} v \sin\delta \sin\varepsilon\right).$$

The numbers of encounters in time dt with distance between p and $p + dp$ and plane making angle between ε and $\varepsilon + d\varepsilon$ is

$$\delta \mathrm{N}\, v\, dt\, p\, dp\, d\varepsilon. \tag{2.3}$$

The mean momentum gained by the particle (1) in time dt by encounters for which $p < \mathrm{P}$ is thus

$$\begin{aligned}\int_{p<\mathrm{P}} \frac{m_1 m_2}{m_1 + m_2} v (1 - \cos\delta)\, p\, dp\, d\varepsilon \,.\, \delta \mathrm{N}\, v\, dt \\ = 4\pi v^2 \frac{m_1 m_2}{m_1 + m_2} \delta \mathrm{N}\, dt \int_0^{\mathrm{P}} \frac{p\, dp}{1 + p^2 v^4/\mu^2} \\ = 2\pi\, (\mu\lambda/v^2)\, \delta \mathrm{N}\, dt \log\left(1 + \mathrm{P}^2 v^4/\mu^2\right) \\ \sim 2\pi\, (\mu\lambda/v^2)\, \delta \mathrm{N}\, dt \log\left(\mathrm{P}^2 v^4/\mu^2\right)\end{aligned} \tag{2.4}$$

in the ξ direction, zero in other directions.

The mean square momentum gained in the ξ-direction is

$$\begin{aligned}\int_{p<\mathrm{P}} \left\{\frac{m_1 m_2}{m_1 + m_2} v (1 - \cos\delta)\right\}^2 p\, dp\, d\varepsilon \,.\, \delta \mathrm{N}\, v\, dt \\ = 8\pi v^3 \left(\frac{m_1 m_2}{m_1 + m_2}\right)^2 \delta \mathrm{N}\, dt \int_0^{\mathrm{P}} \frac{p\, dp}{(1 + p^2 v^4/\mu^2)^2} \\ = 4\pi\, (\lambda^2/v)\, \delta \mathrm{N}\, dt/(1 + \mu^2/\mathrm{P}^2 v^4) \\ \sim 4\pi\, (\lambda^2/v)\, \delta \mathrm{N}\, dt\end{aligned} \tag{2.5}$$

and in either the η or the ζ-direction,

$$\begin{aligned}\int_{p<\mathrm{P}} &\left\{\frac{m_1 m_2}{m_1+m_2}\, v \sin\delta\cos\varepsilon\right\}^2 p\,dp\,d\varepsilon\,.\,\delta\mathrm{N}\,v\,dt\\ &= 4\pi v^3\left(\frac{m_1 m_2}{m_1+m_2}\right)^2 \delta\mathrm{N}\,dt\int_0^{\mathrm{P}}\left\{\frac{1}{1+p^2v^4/\mu^2}-\frac{1}{(1+p^2v^4/\mu^2)^2}\right\}p\,dp\\ &= 2\pi\,(\lambda^2/v)\,\delta\mathrm{N}\,dt\,\{\log\,(1+\mathrm{P}^2v^4/\mu^2)-1/(1+\mu^2/\mathrm{P}^2v^4)\}\\ &\sim 2\pi\,(\lambda^2/v)\,\delta\mathrm{N}\,dt\,\{\log\,(\mathrm{P}^2v^4/\mu^2)-1\}.\end{aligned}\tag{2.6}$$

These formulæ suppose that, while P is so small that the velocities of either particle are unlikely to change much during the encounter owing to the attractions of other particles (condition (3.7) below),

$$\mathrm{P} \gg \mu/v^2. \tag{2.7}$$

3. *Distant Encounters.*

Suppose that in time $d\sigma$ a fraction $f d\sigma$ of the encounters under consideration are terminated, so that v/f is the mean free path for that type of encounter. Then the number of encounters with distance between p and $p+dp$ and plane making angle between ε and $\varepsilon+d\varepsilon$ which start in time dt at an epoch between τ and $\tau+d\tau$ from the time of closest approach and last a time between σ and $\sigma+d\sigma$ is

$$d\mathrm{N}\,v\,dt\,p\,dp\,d\varepsilon\,f^2e^{-f\sigma}\,d\sigma\,d\tau, \tag{3.1}$$

If now $\mathbf{r}$ is vector distance in the relative orbit, $\mathbf{p}$ initial distance of the line of relative motion of one particle from the other, $\mathbf{v}$ initial relative velocity, and t and τ epoch and initial epoch in the relative orbit measured from the time of closest approach, so that $(\mathbf{p}\,.\,\mathbf{v})=0$, then, approximately,

$$\mathbf{r} = \mu\frac{\mathbf{p}+\mathbf{v}t}{(p^2+v^2t^2)^{3/2}}, \qquad \mathbf{r} = \mathbf{v}+\mu\frac{\mathbf{p}t/p^2-\mathbf{v}/v^2}{(p^2+v^2t^2)^{1/2}}-\mu\frac{\mathbf{p}\tau/p^2-\mathbf{v}/v^2}{(p^2+v^2\tau^2)^{1/2}}, \tag{3.2}$$

$$\begin{aligned}\mathbf{r} = \mathbf{p}+\mathbf{v}t &+ \mu\left\{\frac{\mathbf{p}}{p^2v^2}(p^2+v^2t^2)^{1/2}-\frac{\mathbf{v}}{v^3}\log\left(\frac{vt+(p^2+v^2t^2)^{1/2}}{p}\right)\right\}\\ &-\mu t\frac{\mathbf{p}\tau/p^2-\mathbf{v}/v^2}{(p^2+v^2\tau^2)^{1/2}}-\mu\left\{\frac{\mathbf{p}}{p^2v^2}(p^2+v^2\tau^2)^{1/2}-\frac{\mathbf{v}}{v^3}\log\left(\frac{v\tau+(p^2+v^2\tau^2)^{1/2}}{p}\right)\right\}\\ &+\mu\tau\frac{\mathbf{p}\tau/p^2-\mathbf{v}/v^2}{(p^2+v^2\tau^2)^{1/2}},\end{aligned}$$

$$\begin{aligned}\frac{1}{r} = \frac{1}{(p^2+v^2t^2)^{1/2}} &- \mu\left\{\frac{(p^2+v^2t^2)^{1/2}-(p^2+v^2\tau^2)^{1/2}}{v^2\,(p^2+v^2t^2)^{3/2}}\right.\\ &\left.+\frac{(t-\tau)^2}{(p^2+v^2t^2)^{3/2}\,(p^2+v^2\tau^2)^{1/2}}+\frac{t}{v\,(p^2+v^2t^2)^{3/2}}\log\left(\frac{v\tau+(p^2+v^2\tau^2)^{1/2}}{vt+(p^2+v^2t^2)^{1/2}}\right)\right\}.\end{aligned}\tag{3.3}$$

The mean square momentum gained by the particle (1) in the ξ-direction in time dt by encounters for which $p > \mathrm{P}$ is thus

$$\int_{p>\mathrm{P}} \frac{\lambda^2}{v^2}\left\{\frac{1}{(p^2+v^2(\tau+\sigma)^2)^{1/2}} - \frac{1}{(p^2+v^2\tau^2)^{1/2}}\right\}^2 f^2 e^{-f\sigma}\, p\, dp\, d\varepsilon\, d\sigma\, d\tau\,.\,\delta\mathrm{N}\, v\, dt$$

$$= 2\pi\,\frac{\lambda^2}{v}\,\delta\mathrm{N}\, dt\int_{\mathrm{P}}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\left\{\frac{1}{(p^2+v^2(\tau+\sigma)^2)^{1/2}}\right.$$

$$\left. - \frac{1}{(p^2+v^2\tau^2)^{1/2}}\right\}^2 f^2 e^{-f\sigma}\, p\, d\sigma\, d\tau\, dp$$

$$= 2\pi\,\frac{\lambda^2}{v}\,\delta\mathrm{N}\, dt\int_{\frac{f\mathrm{P}}{v}}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\left\{\frac{1}{(l^2+(m+n)^2)^{1/2}} - \frac{1}{(l^2+m^2)^{1/2}}\right\}^2 e^{-n}\, l\, dn\, dm\, dl$$

$$\sim 4\pi\,(\lambda^2/v)\,\delta\mathrm{N}\, dt,^* \tag{3.4}$$

and in either the η or the ζ-direction,

$$\int_{p>\mathrm{P}} \frac{\lambda^2}{p^2}\left\{\frac{\tau+\sigma}{(p^2+v^2(\tau+\sigma)^2)^{1/2}} - \frac{\tau}{(p^2+v^2\tau^2)^{1/2}}\right\}^2 \cos^2\varepsilon\, f^2 e^{-f\sigma}\, p\, dp\, d\varepsilon\, d\sigma\, d\tau\,.\,\delta\mathrm{N}\, v\, dt$$

$$= \pi\,\frac{\lambda^2}{v}\,\delta\mathrm{N}\, dt\int_{\frac{f\mathrm{P}}{v}}^{\infty}\int_{-\infty}^{\infty}\int_0^{\infty}\left\{\frac{m+n}{(l^2+(m+n)^2)^{1/2}} - \frac{m}{(l^2+m^2)^{1/2}}\right\}^2 e^{-n}\,\frac{1}{l}\, dn\, dm\, dl$$

$$\sim 2\pi\,(\lambda^2/v)\,\delta\mathrm{N}\, dt\,\{2\log(2v/f\mathrm{P}) - 2\gamma - 1\}\dagger \tag{3.5}$$

where γ is Euler's constant.

When the above expression for the mean momentum gained in the ξ-direction is integrated in the same way, the result is zero, and comparison with the result for close encounters, (2.4), shows that the term of next higher order in μ or λ is required. This need not be calculated directly. If u is the relative velocity after the encounter, making angle δ with that before (v),

$$-2v\,(u\cos\delta - v) \equiv (u\sin\delta)^2 + (u\cos\delta - v)^2 + (v^2 - u^2).$$

The left-hand side is proportional to the integrand we require, the first two terms on the right-hand side are proportional to the integrands for mean square momentum already integrated. The last term can be treated as follows.

$$\int_{p>\mathrm{P}} (v^2-u^2)\, f^2 e^{-f\sigma}\, p\, dp\, d\varepsilon\, d\sigma\, d\tau\,.\,\delta\mathrm{N}\, v\, dt$$

$$= \int_{p>\mathrm{P}}\left(\frac{2\mu}{r} - \frac{2\mu}{r_0}\right) f^2 e^{-f\sigma}\, p\, dp\, d\varepsilon\, d\sigma\, d\tau\,.\,\delta\mathrm{N}\, v\, dt,$$

---

* See Appendix (7.12).

† See Appendix (7.31).

where r_0 is the initial distance apart. Substituting for $(1/r)$ from (3.3) and retaining the term in μ^2 we obtain,

$$-4\pi \frac{\mu^2}{v} \delta\mathrm{N}\, dt \int_{\mathrm{P}/l}^{\infty} \int_{-\infty}^{\infty} \int_{0}^{\infty} \mathrm{I} e^{-n}\, l\, dn\, dm\, dl,$$

where

$$\mathrm{I} = \left\{ \frac{(l^2 + (m+n)^2)^{1/2} - (l^2 + m^2)^{1/2}}{(l^2 + (m+n)^2)^{3/2}} + \frac{n^2}{(l^2 + (m+n)^2)^{3/2} (l^2 + m^2)^{1/2}} \right.$$
$$\left. + \frac{l+m}{(l^2 + (m+n)^2)^{3/2}} \log \left(\frac{m + (l^2 + m^2)^{1/2}}{m + n + (l^2 + (m+n)^2)^{1/2}} \right) \right\},$$

which tends to

$$-8\pi (\mu^2/v)\, \delta\mathrm{N}\, dt\ *$$

Thus

$$\int_{p>\mathrm{P}} \{-2v(u \cos \delta - v)\} f^2 e^{-f\epsilon} p\, dp\, d\epsilon\, d\sigma\, d\tau \,.\, \delta\mathrm{N}\, v\, dt$$
$$\sim 4\pi (\mu^2/v)\, \delta\mathrm{N}\, dt \{2 \log (2v/f\mathrm{P}) - 2\gamma - 2\},$$

and the mean momentum gained by the particle (1) in time dt

$$\sim 2\pi (\mu\lambda/v^2)\, \delta\mathrm{N}\, dt\, \{2 \log (2v/f\mathrm{P}) - 2\gamma - 2\}. \tag{3.6}$$

These formulæ suppose that while P is so large that the change in velocity in any one encounter is small (condition (2.7) above),

$$\mathrm{P} \ll v/f. \tag{3.7}$$

4. *Formulæ including all Encounters.*

If

$$\mu/v^2 \ll v/f, \tag{4.1}$$

the formulæ of the last two paragraphs can be combined, and we obtain for the mean momentum in the ξ-direction gained by the particle (1) in time dt,

$$4\pi (\mu\lambda/v^2)\, \delta\mathrm{N}\, dt\, \{\log (2v^3/f\mu) - \gamma - 1\}, \tag{4.21}$$

and for the mean square monentum, in the ξ-direction,

$$8\pi (\lambda^2/v)\, \delta\mathrm{N}\, dt, \tag{4.22}$$

and in the η and ζ-directions,

$$4\pi (\lambda^2/v)\, \delta\mathrm{N}\, dt\, \{\log (2v^3/f\mu) - \gamma - 1\}. \tag{4.23}$$

These formulæ can now be summed over particles (2) with various velocities and the variation in v during this summation can be neglected in the argument of the logarithm without great loss of accuracy. In this way, for instance, if the particle (1) has velocity v_1 in the x-direction, and the particles (2) each velocity v_2 in random directions, we find that,

---

* See Appendix (7.21).

$$(a)\ \text{if } v_1 > v_2$$

the mean momentum in the x-direction *lost* by the particle (1) in time dt is

$$4\pi\,(\mu\lambda/v_1^2)\,\delta\mathrm{N}\,dt\,\{\log(2v^3/f\mu) - \gamma - 1\} \tag{4.31}$$

the mean square momentum in the x-direction is

$$\tfrac{2}{3}\,(v_2^2/v_1^3)\,4\pi\lambda^2\,\delta\mathrm{N}\,dt\,\{\log(2v^3/f\mu) - \gamma - 1\} + ((1/v_1) - \tfrac{2}{3}\,(v_2^2/v_1^3))\,8\pi\lambda^2\,\delta\mathrm{N}\,dt, \tag{4.32}$$

and in either perpendicular direction

$$((1/v_1) - \tfrac{1}{3}\,(v_2^2/v_1^3))\,4\pi\lambda^2\,\delta\mathrm{N}\,dt\,\{\log(2v^3/f\mu) - \gamma - 1\} + \tfrac{1}{3}\,(v_2^2/v_1^3)\,8\pi\lambda^2\,\delta\mathrm{N}\,dt, \tag{4.33}$$

while (b) if $v_1 < v_2$,

these quantities are

$$0, \tag{4.41}$$

$$\tfrac{2}{3}\,(1/v_2)\,4\pi\lambda^2\,\delta\mathrm{N}\,dt\,\{\log(2v^3/f\mu) - \gamma - 1\} + \tfrac{1}{3}\,(1/v_2)\,8\pi\lambda^2\,\delta\mathrm{N}\,dt, \tag{4.42}$$

and

$$\tfrac{2}{3}\,(1/v_2)\,4\pi\lambda^2\,\delta\mathrm{N}\,dt\,\{\log(2v^3/f\mu) - \gamma - 1\} + \tfrac{1}{3}\,(1/v_2)\,8\pi\lambda^2\,\delta\mathrm{N}\,dt. \tag{4.43}$$

where v is to be replaced by some mean between $v_2 + v_1$ and $|v_2 - v_1|$.

5. *The Mean Free Path.*

It is apparent from the above formulæ that there is no definite mean free path, but the formulæ are only valid when $v^3 \gg f\mu$ so that considerable uncertainty in f affects them little; a sufficiently close approximation is obtained by putting

$$f = f_1 + f_2 \tag{5.1}$$

where $m_1^2 v_1^2 f_1\,dt$ is equated to the mean square momentum in any direction gained in time dt by particle (1).

For N particles of mean mass m and velocity v per unit volume a rough mean value for f is given by

$$(v^3/\mu f)\log(v^3/\mu f) = 3v^6/4\pi\mu^2\mathrm{N},$$

so that the condition of validity (4.1) takes the form

$$v^6 > \mu^2\mathrm{N}, \tag{5.2}$$

and

$$f = (4\pi\mu^2\mathrm{N}/3v^3)\log(3v^6/4\pi\mu^2\mathrm{N}), \tag{5.3}$$

so that the mean free path is

$$3v^4 \Big/ 4\pi\mu^2\mathrm{N}\log(3v^6/4\pi\mu^2\mathrm{N}). \tag{5.4}$$

In a mixture in nearly equal numbers of electrons and more massive but more slowly moving positive ions, the effects of electrons and ions are of the same order of magnitude in determining the value of f for the electrons, but that of the ions greater than that of the electrons inversely as their velocities in determining the value of f for the ions. If electrons and ions have the same kinetic energy they have nearly the same mean free path.

The correction to Jean's formula (287·4, *loc. cit.*) is to replace in the argument of the logarithm $\nu^{-1/3}$ (ν is our N) by $3v^4/4\pi\mu^2$ N, *i.e.*, to treble the expression; but this would include the effects of stars outside the galaxy where the density is negligible, and the actual result would be, approximately, to double the expression.

6. *The Possible Application to Langmuir's Electrons*

If a beam of 50-volt electrons passes through a cloud of charged particles moving with energy corresponding to, say, 1 volt, the bulk of their effect in scattering the electrons is due to the summation of the effects of many very slight encounters; in most of these the transfer of momentum will be smaller than the momentum of the particles, and there will exist the appropriate probability of the electrons gaining forward momentum. The whole effect must be to set up a distribution with mean change of momentum and mean square change of momentum as calculated from the above formulæ; this distribution would not be a normal probability distribution because there is considerable correlation between the loss of forward momentum and the sideways scattering, but the distribution, say, of forward velocity must be nearly normal. That the gain of random energy given by the above formulæ is of the same order of magnitude as the corresponding loss of forward energy depends essentially on the chance in any one encounter of a gain of forward momentum differing from that of a loss by a quantity much smaller than either (see the deduction of (3.6) above). If this were not so, if for instance only loss of energy were possible, the gain of random energy would be much smaller than the loss of forward energy, for the latter is proportional to the product of the momentum and the mean loss of momentum, the former to the mean square loss of momentum.

The formula giving the loss of momentum in distance dx from an electron of charge e, mass m_1, and velocity v_1, to slower electrons or ions of charge e_2, mass m_2, and velocity v_2, distributed at N per unit volume is (4.31)

$$4\pi \frac{m_1 + m_2}{m_1 m_2} \frac{e^2 e_2^2}{v_1^3} \mathrm{N}\, dx \left\{ \log \frac{2 v_1^3 m_1 m_2}{f(m_1 + m_2) e e_2} - \gamma - 1 \right\}. \qquad (6.1)$$

The corresponding formula for the longitudinal mean square momentum change is (4.32)

$$8\pi \frac{e^2 e_2^2}{v_1^2} \mathrm{N}\, dx \left[1 + \frac{1}{3} \frac{v_2^2}{v_1^2} \left\{ \log \frac{2 v_1^3 m_1 m_2}{f(m_1 + m_2)\, e e_2} - \gamma - 3 \right\} \right] \tag{6.2}$$

If $\mathrm{N} = 4{\cdot}6\ \ 10^{13}$, the value of f for doubly charged mercury ions at "temperature 3500°," *i.e.*, with mean energy $7{\cdot}2 \,.\, 10^{-13}$ ergs., is, by (5.3), $2{\cdot}5 \,.\, 10^7$, corresponding to a mean free path $1{\cdot}5 \,.\, 10^{-3}$ cm.; for electrons at temperature 95,000°, it is $6{\cdot}4 \,.\, 10^7$ corresponding to mean free path $3{\cdot}2$ cm., for electrons at temperature 15,600°, $7{\cdot}3 \,.\, 10^8$ corresponding to mean free path $0{\cdot}11$ cm. In these three cases, if the primary electron velocity (v_1) corresponds to 50 volts, $\log \{2 v_1^3 m_1 m_2 / f(m_1 + m_2) e e_2\}$ has the values respectively $23{\cdot}16$, $22{\cdot}24$, $19{\cdot}80$; a tenfold decrease in N adds $2{\cdot}3$ to these, a tenfold decrease in the "temperature" subtracts $3{\cdot}4$, and a tenfold decrease in the primary electron energy also subtracts $3{\cdot}4$.

If we define the increase of random (longitudinal) energy $\Delta \mathrm{E}_r$ as $\frac{1}{2} k\mathrm{T}$ where T is temperature corresponding to the distribution of longitudinal velocity of the primary beam about its mean (Langmuir, *loc cit*, p 607), and if the mean decrease of (longitudinal) energy (*i e*, the decrease of energy of a particle with the mean velocity) is $\Delta \mathrm{E}_l$, we have from (6.1) and (6.2) for the effect of ions,

$$\frac{\Delta \mathrm{E}_l}{\Delta \mathrm{E}_r} = 4\pi \left\{ \log \frac{2 v_1^3 m_1 m_2}{f(m_1 + m_2)\, e e_2} - \gamma - 1 \right\} \Big/ \frac{1}{2} 8\pi,$$

(since the term in v_2^2/v_1^2 will be negligible), while for electrons

$$\frac{\Delta \mathrm{E}_l}{\Delta \mathrm{E}_r} = 8\pi \left\{ \log \frac{2 v_1^3 m_1 m_2}{f(m_1 + m_2)\, e e_2} - \gamma - 1 \right\} \Big/ \frac{1}{2} 8\pi \left[1 + \frac{1}{3} \frac{v_2^2}{v_1^2} \left\{ \log \frac{2 v_1^3 m_1 m_2}{f(m_1 + m_2)\, e_1 e_2} - \gamma - 3 \right\} \right].$$

If, however, the primary beam is emitted from a fine wire, scattering near the wire in a plane perpendicular to the wire will not give the electrons energy transverse to the general motion away from the wire at the expense of their longitudinal energy but will only transfer them to another sector of the beam. Thus if most of the scattering takes place near the wire the transverse energy gained by scattering perpendicular to the wire (obtained from (4.33)) must be added to the mean longitudinal energy (obtained from (4.31)), *i e.*, subtracted from $\Delta \mathrm{E}_l$. Comparing (4.33), of which only the first term is important, and (4.31) we see that this will reduce $\Delta \mathrm{E}_l$ due to ions by one-half and due to electrons by one-quarter. Thus if the scattering were due to $4{\cdot}6 \,.\, 10^{13}$ electrons, at temperature 95,000°, and the same number of (singly) ionised mercury atoms, at temperature 3500°, per unit volume close to the wire cathode, we should have

$$\frac{\Delta E_l}{\Delta E_r} = \frac{6\ \ 19{\cdot}7 + 2\,.\,22{\cdot}6}{12 + 4\,.\,\frac{1}{4}\,.\,18{\cdot}7 + 12} = 3{\cdot}8.$$

Langmuir (*loc. cit.*, p. 601) has obtained values varying from some such value as this as a minimum when ΔE_r is particularly large to values of the order of 80 for cases where ΔE_r and ΔE_l are both considerably smaller. It does not seem impossible that his observations may be explained as the combination of scattering by electrons and ions which takes place only where there is large primary current density and a smaller effect, due perhaps to encounters with atoms in which energy is much more easily lost than gained and which would give larger values of $\Delta E_l/\Delta E_r$.

The very large values of ΔE_r would require large concentrations of ions and secondary electrons to produce them. To produce a longitudinal temperature 56,000° in a beam of 50 volt electrons would, by (6.1) require secondary electrons with energy corresponding to temperature 15,600° and singly-ionised mercury atoms with a value of $N\,dx$ of $4{\cdot}2\,.\,10^{14}$, or doubly-ionised mercury atoms with two electrons at 95,000° each with a value of $N\,dx$ of $1{\cdot}05\,.\,10^{14}$. In Langmuir's experiment in which the longitudinal temperature produced was 56,000° (*loc. cit.*, p 599), the number of mercury atoms present was $4{\cdot}6\,.\,10^{13}$ per cubic centimetre and the distance traversed was 3 cm. To explain the observed value of ΔE_r nearly all the mercury atoms in the path of the beam must be doubly ionised. This seems to be too many, but it must be remembered that if such large concentrations of ions are present where they are being produced they need not be present throughout the tube. If the 0·42 watt lost by the primary beam in the above case went mostly to maintaining $2{\cdot}0\,.\,10^{15}$ 95,000° secondary electrons, these would have to last on the average $9{\cdot}3\,.\,10^{-3}$ seconds. Potential gradients would be set up sufficient to stop electrons diffusing away from where they were produced faster than the positive ions, and these, moving at $1{\cdot}05\,.\,10^{5}$ cm./sec. with a mean free path of $1{\cdot}5\,.\,10^{-2}$ cm. can diffuse in $9{\cdot}3\,.\,10^{-3}$ seconds only about 1·0 cm.

On the whole it seems not entirely impossible that Langmuir's observations can be accounted for as scattering due to positive ions and secondary electrons.

APPENDIX.

$$\int_{\epsilon}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}\left\{\frac{1}{(l^2+(m+n)^2)^{1/2}}-\frac{1}{(l^2+m^2)^{1/2}}\right\}^2 e^{-n}\, l\, dn\, dm\, dl \tag{7.1}$$

$$= \int_{-\infty}^{\infty}\int_{m}^{\infty}\left[\int_{\epsilon}^{\infty}\left\{\frac{1}{(l^2+n^2)^{1/2}}-\frac{1}{(l^2+m^2)^{1/2}}\right\}^2 l\, dl\right] e^{m-n}\, dn\, dm$$

$$= \int_{-\infty}^{\infty}\int_{m}^{\infty} e^{m-n}\log\frac{(\epsilon^2+m^2)^{1/2}(\epsilon^2+n^2)^{1/2}+\epsilon^2+\frac{1}{2}(m^2+n^2)}{2(\epsilon^2+m^2)^{1/2}(\epsilon^2+n^2)^{1/2}}\, dn\, dm. \tag{7.11}$$

The infinity of the integrand of

$$\int_{-\infty}^{\infty}\int_{m}^{\infty} e^{m-n}\log(\epsilon^2+n^2)\, dn\, dm$$

lies at $\epsilon = n = 0$;

$$\left|\int_{-\infty}^{\eta}\int_{-\eta}^{\eta}+\int_{-\eta}^{\eta}\int_{m}^{\eta} e^{m-n}\log(\epsilon^2+n^2)\, dn\, dm\right|$$

$$= \left|\int_{-\eta}^{\eta}\int_{-\infty}^{n} e^{m-n}\log(\epsilon^2+n^2)\, dm\, dn\right|$$

$$= \left|2\eta\log(\epsilon^2+\eta^2)-4\eta+4\epsilon\tan^{-1}\eta/\epsilon\right|$$

$$< 4\eta\,|\log\eta| + 4\eta,$$

provided that $\eta < \frac{1}{2}$, $\epsilon < \frac{1}{2}$ and since $\tan^{-1}\eta/\epsilon < \eta/\epsilon$. Hence

$$\int_{-\infty}^{\infty}\int_{m}^{\infty} e^{m-n}\log(\epsilon^2+n^2)\, dn\, dm$$

converges uniformly in ϵ as $\epsilon \to 0$. Similar assertions hold for the other terms of (7.11). Hence

$$\operatorname*{Lt}_{\epsilon\to 0}\int_{\epsilon}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}\left\{\frac{1}{(l^2+(m+n)^2)^{1/2}}-\frac{1}{(l^2+m^2)^{1/2}}\right\}^2 e^{-n}\, l\, dn\, dm\, dl$$

$$= \int_{-\infty}^{\infty}\int_{m}^{\infty} e^{m-n}\log\frac{(|m|+|n|)^2}{4|mn|}\, dn\, dm$$

$$= 2\int_{0}^{\infty}\int_{m}^{\infty}(e^{m-n}+e^{-m-n})\log\frac{(m+n)^2}{4mn}\, dn\, dm$$

$$= 2\int_{0}^{\infty}\int_{0}^{\infty} e^{-n}\log\frac{(2m+n)^2}{4m(m+n)}\, dm\, dn + 2\int_{0}^{\infty}\int_{0}^{n/2} e^{-n}\log\frac{n^2}{4m(n-m)}\, dm\, dn$$

$$= 2\int_{0}^{\infty} e^{-n}\, n\log 2\, dn + 2\int_{0}^{\infty} e^{-n}\, n(1-\log 2)\, dn$$

$$= 2 \tag{7.12}$$

$$\operatorname*{Lt}_{\epsilon\to 0}\int_{\epsilon}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty} \mathrm{I}e^{-n}\, l\, dn\, dm\, dl \tag{7.2}$$

is calculated in the same way, the integral converging uniformly in ϵ.

$$\int_0^\infty \mathrm{I}l\,dl = 2\log\frac{2m+n}{2(m+n)} + \frac{2n}{2m+n} \qquad m > 0 \quad \text{or} \quad m+n < 0$$

$$= 2 - 2\log 2 \qquad\qquad -n < m < 0$$

$$\int_0^\infty \int_{-\infty}^{-n} + \int_0^\infty \int_0^\infty e^{-n}\left(2\log\frac{2m+n}{2(m+n)} + \frac{2n}{2m+n}\right) dm\,dn = 2\log 2$$

$$\int_0^\infty \int_{-n}^0 e^{-n}(2 - 2\log 2)\,dm\,dn = 2 - 2\log 2,$$

so that

$$\operatorname*{Lt}_{\varepsilon \to 0} \int_\varepsilon^\infty \int_{-\infty}^\infty \int_0^\infty \mathrm{I}e^{-n}\,l\,dn\,dm\,dl = 2, \tag{7.21}$$

$$\int_\varepsilon^\infty \int_{\infty}^\infty \int_0^\infty \left\{\frac{m+n}{(l^2+(m+n)^2)^{1/2}} - \frac{m}{(l^2+m^2)^{1/2}}\right\}^2 e^{-n}\frac{1}{l}\,dn\,dm\,dl \tag{7.3}$$

$$= \int_{-\infty}^\infty \int_m^\infty \left[\int_\varepsilon^\infty \left\{\frac{n}{(l^2+n^2)^{1/2}} - \frac{m}{(l^2+m^2)^{1/2}}\right\}^2 \frac{1}{l}\,dl\right] e^{m-n}\,dn\,dm.$$

$$\int_\varepsilon^\infty \left\{\frac{n}{(l^2+n^2)^{1/2}} - \frac{m}{(l^2+m^2)^{1/2}}\right\}^2 \frac{dl}{l}$$

$$= \log\left[\frac{(m+n)^2}{2mn}\,\frac{(\varepsilon^2+m^2)^{1/2}(\varepsilon^2+n^2)^{1/2}}{(\varepsilon^2+m^2)^{1/2}(\varepsilon^2+n^2)^{1/2}+mn+\varepsilon^2(m^2+n^2)/2mn}\right]$$

$n > m > 0$ or $0 > n > m$

$$= \log\left[\frac{-2mn}{(m-n)^2}\,\frac{(\varepsilon^2+m^2)^{1/2}(\varepsilon^2+n^2)^{1/2}\{(\varepsilon^2+m^2)^{1/2}(\varepsilon^2+n^2)^{1/2}-mn-\varepsilon^2(m^2+n^2)/2mn\}}{\varepsilon^4}\right]$$

$n > 0 > m.$

All the terms converge uniformly in ε except that corresponding to ε^4 in the denominator of this last expression.

Hence

$$\int_\varepsilon^\infty \int_{-\infty}^\infty \int_0^\infty \left\{\frac{m+n}{(l^2+(m+n)^2)^{1/2}} - \frac{m}{(l^2+m^2)^{1/2}}\right\}^2 e^{-n}\frac{1}{l}\,dn\,dm\,dl$$

$$\underset{\varepsilon\to 0}{\sim} 2\int_0^\infty \int_0^\infty e^{-n}\log\frac{(2m+n)^2}{4m(m+n)}\,dm\,dn + 2\int_0^\infty \int_0^{n/2} e^{-n}\log\frac{4m^3(n-m)^3}{\varepsilon^4 n^2}\,dm\,dn$$

$$= 2\int_0^\infty e^{-n}n\log 2\,dn + 2\int_0^\infty e^{-n}n\{2\log n - 3 + \log 2 - 2\log\varepsilon\}\,dn$$

$$= 4\log(2/\varepsilon) - 4\gamma - 2. \tag{7.31}$$

The Refractive Index of Quartz.

By Dr. W. R. C. Coode Adams.

(Communicated by T. M. Lowry, F R S.—Received August 21, 1928)

In the previous paper* an equation was put forward connecting the wave-length and the refractive index of quartz for the ordinary ray. It is the purpose of this paper to put forward a similar equation for the extraordinary ray.

The equation adopted was of the Kettler-Helmholtz type

$$n^2 = n_\infty{}^2 + \Sigma \frac{M_m}{\lambda^2 - \lambda_m{}^2},$$

where λ_m is in each case the wave-length corresponding to the natural frequency of an electron and M_m the corresponding refraction constant.

The values of λ_m were obtained from some previous work on optical rotatory power of quartz.† The other constants were solved from simultaneous equations arising from putting in known values for the refractive index for known wave-lengths.

In this present paper the same values for the natural frequencies were used and the same refraction constant for the infra-red as it was shown that this had very little effect. Values for the refractive index of the extraordinary ray for known wave-lengths were taken and the other constants solved from the simultaneous equations resulting.

The data used, as in the previous case, was that contained in a paper by Gifford‡ with certain corrections for recent measurements of wave-lengths.

The following equation resulted where n is the refractive index and λ the wave-length:

$$n^2 = 3\cdot5612557 + \frac{0\cdot00844614}{\lambda^2 - 0\cdot0127493} + \frac{0\cdot00276113}{\lambda^2 - 0\cdot000974} + \frac{127\cdot2}{\lambda^2 - 108}.$$

Below is given a table showing the calculated and observed refractive indices for 18 lines in the visible and ultra-violet.

As most of the measurements of refractive index in the infra-red seem to have been done with the ordinary ray at present it is not possible to extend the research further in that direction.

---

* 'Roy. Soc. Proc.,' A, vol. 117, p. 209 (1927).

† Lowry and Coode Adams, 'Phil. Trans.,' A, vol. 226, pp. 391–466 (1927).

‡ 'Roy. Soc. Proc.,' vol. 70, p. 329 (1902).

Extraordinary Ray.

Wave-length.	Observed	Calculated.	Difference.
7950·46 Rb	1·54742	1·54744	−0·00002
7685·23 K	1·54800	1·54800	±
7065·20 He	1·54949	1·54952	−0·00003
6563·04 H	1·55095	1·55095	±
5892·948 Na	1·55337	1·55335	+0·00002
5270·36 Fe	1·55639	1·55638	+0 00001
4861·49 H	1·55899	1·55902	−0·00003
4340·66 H	1·56341	1·56344	−0·00003
3961·68 Al	1·56783	1·56782	+0 00001
3610·66 Cd	1·57324	1·57329	−0·00005
3302·80 Zn	1·57973	1·57971	+0·00002
3034·21 Sn	1·58723	1·58722	+0·00001
2748·71 Cd	1·59813	1·59817	−0·00004
2573·10 Cd	1·60716	1·60719	−0·00003
2312·98 Cd	1·62565	1·62564	+0·00001
2265·11 Cd	1·62996	1·62991	+0·00005
2194·60 Cd	1·63702	1·63706	−0·00004
2144·35 Cd	1·64270	1·64270	±

The Mode of Formation of Neumann Bands. Part I.—*The Mechanism of Twinning in the Body-Centred Cubic Lattice.*

By S. W. J. Smith, F.R.S., A. A. Dee, B.Sc., D.I.C., and J. Young, B.Sc., F.R.A.S., Physical Laboratory, University of Birmingham.

(Received September 24, 1928.)

1. The existence of Neumann lamellæ as a characteristic feature of hexahedral meteoric iron and of the kamacite of octahedral meteorites has been known for many years. The traces of these lamellæ upon polished and etched surfaces were at first regarded as Widmanstätten figures; but it was shown by Neumann that they were distinct from such figures, that they were characteristic of single cubic crystals, and that their outcrops upon a cube face, which he determined, were inconsistent with the assumption that they were octahedral lamellæ.

Neumann, Tschermak and other mineralogists inferred from the geometrical relations between the outcrops of the bands that they were a consequence of an interpenetrating-cube twin structure within the meteorite, of a type known to occur in various minerals, *e.g.*, in fluor spar.

Subsequently Linck[*] suggested what is now generally assumed to be the correct description of the orientation of the lamellæ with respect to the matrix. His observations were made upon a cleavage fragment of artificial iron from which, by appropriate etching and goniometric measurement, he determined the orientations of the lamellæ with respect to the cube faces and concluded that they occupied the positions of icositetrahedral {112} planes.

The same conclusion was reached more directly by Osmond and Cartaud[†] who cut a specimen with cube faces from a large crystal of artificial iron and traced the outcrops of individual lamellæ upon these faces.

Although the positions of the lamellæ were thus determined, no direct evidence was obtained concerning the crystallographic orientation of the material in them with respect to that in the matrix. Linck assumed that, as in the case of many minerals, the matrix and the lamellæ were in twin relation with respect to the {112} planes, and that the material did not consist of interpenetrant twins as had been at first supposed. He suggested that the twinning was probably due to slipping and might result from such causes as fracture or change of temperature.

As evidence of movement during the formation of the lamellæ he pointed out that when, as frequently happened, thin rhabdite crystals were present in the matrix, they were sometimes broken where they were crossed by Neumann bands. But, although such observations do suggest that there is some kind of slip when Neumann lamellæ are produced, they do not prove that twin formation occurs at the same time and, indeed, in a recent paper[‡], it has been argued that twinning of the kind postulated by Linck does not occur. Other methods of interpreting Neumann bands have also been suggested.[§]

In the course of an examination of the magnetic and other properties of meteoric iron we have had occasion to consider again the problem presented by Neumann lamellæ in materials of the kind in which they were first observed. We hope to show beyond doubt, in what follows, not only that Neumann bands are due to twinning but also exactly how they arise.

2. Artificial or mechanical twinning may be regarded as the result of a movement within a crystal by which the orientation of the atoms, in a band

---

[*] 'Z. Krystallog.,' vol. 20, p. 209 (1892), see also, Cohen, 'Meteoritenkunde,' vol. 1, pp. 89 *et seq.* (1894).

[†] 'J. Iron & Steel Inst.,' vol. 71, pp. 444 *et seq.* (1906).

[‡] Rosenhain and McMinn, 'Roy. Soc. Proc.,' A, vol. 108, p. 235 (1925).

[§] *Cf.*, *e.g.*, Thompson and Millington, 'J. Iron & Steel Inst.,' vol. 109, p. 67 (1924).

bounded by parallel planes, becomes a mirror image with respect to these planes of that in the unchanged matrix on either side.

The effect of the movement may be represented as in fig. 1. Here S_1 and S_2 represent planes perpendicular to the plane of the paper, which form the boundaries of the band and which are parallel to a face of the form $\{hkl\}$ of the system to which the matrix conforms.

The equidistant parallel lines, to the left of S_1, represent a series of planes parallel to a face of the form $\{h'k'l'\}$ in which the atoms of the matrix can be supposed to be distributed. The material is to be regarded as so oriented that the lines of intersection of these planes with the plane S_1 are at right angles to the plane of the paper, *i.e.*, the atom planes, like the twin planes, are perpendicular to the plane of the paper.

Between S_1 and S_2, the atom planes of the twinned matrix are mirror images in S_1 of the corresponding atom planes to the left of S_1 and are similarly mirror images in S_2 of the atom planes to the right of S_2.

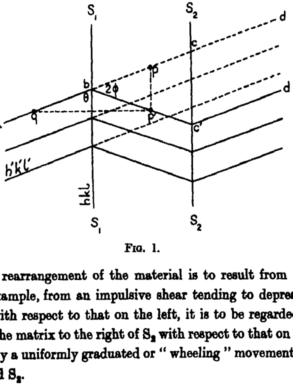


FIG. 1.

Now if this rearrangement of the material is to result from a mechanical impulse, for example, from an impulsive shear tending to depress the matrix on the right with respect to that on the left, it is to be regarded as a bodily movement of the matrix to the right of S_2 with respect to that on the left of S_1, accompanied by a uniformly graduated or " wheeling " movement of the atoms between S_1 and S_2.

The problem is to examine, for a given lattice, the different ways in which a movement of the kind pictured in fig. 1 is possible, and to decide which of them is likely to be produced most easily.

Taking any atom plane $a\ b\ c\ d$, from which the new configuration $a\ b\ c'\ d'$ arises, any atom p, between S_1 and S_2, moves into a position p' such that $bp = bp'$ and that p' is a mirror image of an atom q. Consequently $bp' = bq = bp$, where p, p' and q are either in the plane of the paper or in a parallel plane above or below it. Hence the atoms in the plane $a\ b\ c\ d$ must be so distributed that, along any line in this plane parallel to the plane of the paper, they lie symmetrically about the point in which the line cuts the plane S_1. In other words, the line in which the plane $a\ b\ c\ d$ cuts S_1 must be an axis of symmetry for the atomic distribution in the plane. The obvious places, therefore, in which to look for simple twinning conditions, are those in which the plane of fig. 1 is a plane of symmetry of the matrix.

3. We have found by X-ray analysis that the hexahedral meteorite (Coahuila), upon which most of our experiments were made, has a body centred lattice closely similar to that of α-iron, and that its parameter is $2{\cdot}88 \,.\, 10^{-8}$ cm. The constituent—kamacite—of the octahedral meteorites, to which the bands are confined, is also body centred and of practically the same parameter. The conditions of twinning to be examined are therefore those of the body centred lattice.

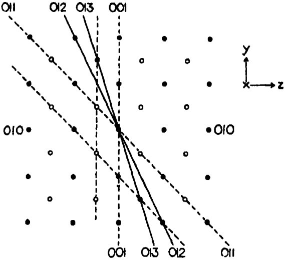


Fig. 2.

Fig. 2 shows the distribution of the atoms of this lattice with respect to a cube plane (100), parallel to the plane of the paper.

The dots indicate the centres of atoms in the plane of the paper and the circles those of atoms in the adjacent parallel plane below it. The distribution of the atoms in succeeding parallel planes are repetitions of those occurring in these two, each dot and each circle shown being at the end of a row of atoms, uniformly spaced, extending downwards at right angles to the paper.

This figure, considered in conjunction with fig. 1, shows that planes parallel to 012 and to 013, respectively, are possible twin planes corresponding with *hkl* of that figure. Similarly, atom planes parallel to 001 and to 011, respectively, indicated by dotted lines, contain for each of the twin planes, the distribution of atoms required in *h'k'l'* of the same figure.

From fig. 1 the angular wheeling movement during twinning, *i.e.*, the angle between *bc* and *bc'*, is determined by the angle between the twin plane *hkl* and the atom plane *h'k'l'*. Denoting the latter angle by θ and the former by 2ϕ, the relation between the two angles is $\phi = \frac{1}{2}\pi - \theta$.

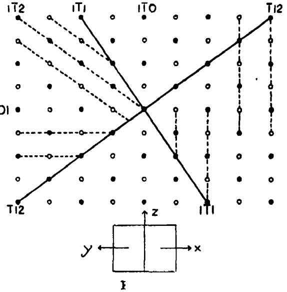
Consequently, the larger the (acute) angle between the possible twin plane and the possible atom plane, the smaller is the angular movement required to produce the twin orientation of the band with respect to the matrix. The values of θ and ϕ for the different cases are tabulated below :—

hkl.	*h'k'l'.*	θ.		ϕ.		Tan ϕ.	Cos ϕ.
		°	′	°	′		
012	001	26	34	63	26	2	0 448
013	011	26	34	63	26	2	0 448
012	011	18	26	71	34	3	0 318
013	001	18	26	71	34	3	0·318

The reason for the tabulation of cos ϕ will appear later when the details of the atomic movement during twinning are discussed. Even in the most favourable of these cases ($\phi = 63° 26'$), the angular movement required to give twinning is very large.

4. Fig. 3 shows the distribution of the atoms in the body centred lattice with respect to a rhombic dodecahedral plane (110) parallel to the plane of the paper, the dots and circles having respectively the same significance as in fig. 2.

This figure, in conjunction with fig. 1, shows that planes parallel to $\bar{1}12$ and to $1\bar{1}1$, respectively, are possible twin planes. With respect to the former, atom planes parallel to $1\bar{1}0$, $1\bar{1}2$ and 001, respectively, indicated by dotted lines, contain the distributions of atoms required of *h'k'l'* in fig. 1. With respect to $1\bar{1}1$, atom planes parallel to $1\bar{1}0$ and to $1\bar{1}2$, respectively, satisfy



the required condition. The values of θ and ϕ for the different cases are :—

hkl.	$h'k'l'$.	θ.	ϕ.	Tan ϕ.	Cos ϕ.
		° ′	° ′		
$\bar{1}12$	$1\bar{1}2$	70 32	19 28	$1/2\sqrt{2}$	0·943
$\bar{1}12$	$1\bar{1}0$	54 44	35 16	$1/\sqrt{2}$	0·816
$\bar{1}12$	001	35 16	54 44	$\sqrt{2}$	0 577
$1\bar{1}1$	$1\bar{1}0$	35 16	54 44	$\sqrt{2}$	0·577
$1\bar{1}1$	$1\bar{1}2$	19 28	70 32	$2\sqrt{2}$	0·333

Inspection of the data in this and in the preceding table leaves no doubt as to which of the possible forms of twinning is likeliest to occur. It is that of the type $hkl = \bar{1}12$; $h'k'l' = 1\bar{1}2$. By symmetry, the same type of twinning will occur with $1\bar{1}2$ as twin plane and $\bar{1}12$ parallel to the atom planes.

To each of the six planes of the form {110} two of the twelve planes of the form {112} are perpendicular. Accordingly there are twelve different directions in which twinning of the "easiest" type can occur within the matrix.

5. Fig. 4 indicates the essential features of this type of twinning. A B C D, to the left of the twin plane SS, shows the orientation of an element of the body centred lattice. In this, as is well known, each atom can be regarded as in contact with eight others, the assumption being that atoms whose centres

are at the minimum distance apart "touch" each other. The atoms are therefore in "contact" in rows parallel to each of the four cube diagonals.

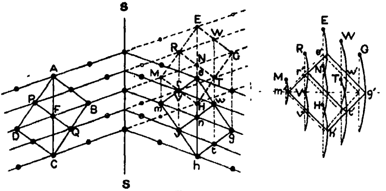


FIG. 4.

In fig. 4 two of these rows, parallel to AC and to BD respectively, are in the (110) plane of the paper. The others, parallel to PFQ, are in the $(1\bar{1}0)$ plane perpendicular to the paper.

The eight atoms with which F is in contact are A, C, B and D in the plane of the paper, P and Q in the first (110) atom plane below the paper and P′ and Q′, vertically above P and Q, in the first (110) atom plane above the paper.

E, G, H, M shows the orientation of an element of the matrix lattice, before twinning occurs, in the region to the right of SS. The figure shows how, as a result of movements of the type shown in fig. 1, this element of the matrix lattice changes into an element *e*, *g*, *h*, *m* of a precisely similar cubic lattice of which, however, the orientation is a mirror image of the matrix lattice with respect to the twin plane.

6. The movements of the atoms of the lattice element as they pass from their initial to their final positions are an essential feature of the twinning process. It is therefore important to attempt to define them as accurately as possible. We assume, with this object, that atoms in line and in contact in any (110) plane like that of fig. 4 remain in line and in contact during the twinning process. With this assumption the tracks of the atoms will lie between the two extremes indicated on the right in fig. 4.

If, during the movement, the spacing of the atoms remains constant, then the atoms M, R, E, W, G, etc., will move in the circular arcs shown. This will imply a temporary increase in the "width" of the band of the matrix in which the twinning occurs. At the point of maximum "expansion," half-way through the twinning process, the lattice element will be of the form indicated by the continuous lines m', r', e', w', g', etc. The atoms M, E, G, H in the (110) plane of the paper will remain in contact with N throughout the twinning movement, although they may turn on their axes during the process. The atoms R, W, T, V in the (110) plane immediately below the paper will remain in contact, each with two of the others, throughout the movement; but the contacts between this set of atoms and those in the plane of the paper will change. At the beginning R and T will be in contact with N, while W and V will not. W will be in contact with E and G and V with M and H.

At all intermediate stages there will be no contacts between the atoms in one (110) plane and the atoms in the neighbouring planes. The "openness" of the structure will be at its maximum in the midway position. Here W will be at the same distance from E and G as R is from E and M; although, at the outset, W, E and G were in contact. As the movement continues R will come nearer to E and M, while the distance between W and E, G will increase. Finally the twin position will be reached with R, E and M in contact and W, E and G at the maximum distance apart.

7. Without any hypothesis as to the nature of the interatomic forces, we can safely assert that under their operation the body centred lattice is a relatively stable configuration. Half-way through the twinning movement, the atoms lie midway between two forms of this configuration. In this position the slightest bias, forward or backward, would tend to make the system move, under the operation of the interatomic forces alone, forward toward the twin lattice or backward toward the original one. The midway position is therefore one of unstable equilibrium.

If the shearing impulse is insufficient to carry the movement beyond this position the interatomic forces, when it has passed, will restore the original orientation.

If the impulse is just sufficient to carry the movement beyond the midway position, the interatomic forces will facilitate the passage to the twin position.

Even when the impulse is considerably more than sufficient to effect the transformation, the material will still come to rest in the twin position; because, in this, it is much more resistant than it was originally to impulses of the kind which initiated the movement.

8. An equally definite picture, of the way in which a shearing impulse of the right direction can produce a mechanical twin, is obtained by assuming that, during the twinning process, the atoms move in lines parallel to the

trigonal axis in which the {112} twin plane cuts the {110} plane of the paper, and not in arcs of circles.

This case is indicated by the dotted lines in the diagram on the right in fig. 4. The spacing of the atoms like M, N and G, in rows parallel to the plane of the paper, will not now remain constant during the twinning process. It will decrease continuously to a minimum at the midway position. Thereafter it will increase towards its original value which it will recover when the twin position is attained. In this case, the interatomic forces will resist the decrease in the atomic spacing. The resistance will be a maximum at the midway position which, as before, will be unstable. Until this position is reached the compressed atoms, left to themselves, will return to their original positions; beyond it, they will move into the twin formation.

The actual movement during twinning probably lies between the two which have been described, and the geometry of the one does not differ greatly from that of the other as fig. 4 shows. This happens because the divergence between the two movements depends upon the difference between cos ϕ and unity, which is much less in this easiest type of twinning than in any of the others, as the tables on pp. 481 and 482 show.

9. Before leaving the question of mechanical twinning, one point of some importance remains to be discussed. The capacity of the matrix to pass from one stable configuration to another, which is its mirror image in a particular plane, has been taken to be an essential feature of slip under impulsive stresses as distinct from stresses steadily applied or steadily increased. It should be pointed out, however, that there must always be some distortion in the immediate neighbourhood of the twin plane. In some cases it would be small and in others very large. An example, of the former type, is seen in fig. 4. Here, in the 110 plane just below the plane of the paper, the pairs of atoms immediately facing each other across the twin plane are nearer together than any atoms were, before the twinning took place, in the ratio 0·943 : 1.

If the twin plane were $1\bar{1}1$, there would be a distortion of similar character but greater magnitude, the distance between the centres of the "compressed" atoms being now only two-thirds of the closest spacing in the undistorted lattice.

Much more pronounced effects of the same kind would occur with 012 as the twin plane, as in fig. 2.

On occasion, this feature of the twinning process might be the deciding factor between one type of twinning and another.

10. From the experimental sections which follow, it will be seen that the

most direct proof of the twin relationship between bands and matrix is afforded by a study of the pits produced simultaneously on both, by suitable etchants. We have seen, however, that the mirror image relation can result from each of several different atomic movements with respect to the same twin plane. Consequently the etching pits, alone, do not suffice to fix the precise character of the movement. For this purpose, the direction and the extent of the movement producing the bands must be determined.

The Mode of Formation of Neumann Bands. Part II.—*The Evidence that the Bands are Twins.*

By S. W. J. Smith, F.R.S., A. A. Dee, B.Sc., D.I.C., and J. Young, B.Sc., F.R.A.S., Physical Laboratory, University of Birmingham.

(Received September 24, 1928.)

[Plates 4–7.]

1. Any discussion of the significance of Neumann bands must involve the geometrical relationships between them and the matrix.

The orientation of the cube lattice of a meteorite having been determined by X-ray or other methods, a section parallel to a cube face can be cut, polished and etched, and the angles which the traces of the Neumann bands make with the sides or the diagonals of the cube face can be measured. If the bands are parallel to the twelve planes of the icositetrahedron {112}, their traces will lie

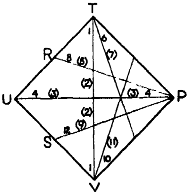


Fig. 5.

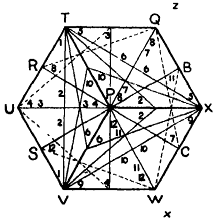


Fig. 6.

in the directions shown in fig. 5, PVUT representing the orientation of the cube face. The spatial relationships of the planes producing such {112} traces can be visualised by means of fig. 6, an isometric projection of a cube, in which the planes producing the traces shown in fig. 5 are indicated by their traces on three mutually perpendicular planes, *e.g.*, PVUT (010), PTQX (001) and PXWV (100).

In these figures, to avoid confusion, the {112} planes are denoted by the numbers 1 to 12 instead of by their indices. The latter can be written :—

1	$1\bar{2}1$	5	$11\bar{2}$	9	$\bar{2}11$
2	121	6	112	10	211
3	$1\bar{2}\bar{1}$	7	$\bar{1}1\bar{2}$	11	$\bar{2}1\bar{1}$
4	$\bar{1}\bar{2}1$	8	$1\bar{1}\bar{2}$	12	$\bar{2}\bar{1}1$

We have examined, in detail, the phenomena presented by the bands on four planes of section, respectively parallel to faces of the matrix 010, 101, 111 and 212, corresponding with the planes PVUT, VTQW, VTX and VTBC of fig. 6.

Photomicrographs of these sections can be similarly orientated by means of their common diad axis TV, to which the traces of the planes 1, 2 are parallel in all of them.

The angles which the traces of the remaining ten planes make with TV can be measured for each plane of section. These angles can also be calculated on the assumption that the twelve planes conform to {112}.

In the meteorites we have examined, the two sets of angles agree within the limits of experimental error. Consequently, in them, the Neumann bands lie in {112} planes.

The calculated angles, to be compared with those determined experimentally, are collected below ; the angles (*a*) which the traces of the other {112} planes make with that of 1, 2 being followed in each case by the angles (*b*) which the {112} planes make with the plane of section :—

Section 010.

a :—(10, 11) 18° 26′ ; (9, 12) 71° 34′ ; (3, 4) 90° ; (6, 7) −18° 26′ ; (5, 8) − 71° 34′.

The sign convention will be obvious from fig. 5.

b :—(1, 2, 3, 4) 35° 16′ ; (5 . . . 12) 65° 54′.

The directions of these angles can be visualised by the aid of fig. 6.

Section 101.

a :—(3, 6, 11) 35° 16′; (8, 9) 64° 45′; (4, 7, 10) — 35° 16′; (5, 12) — 64° 45′.

b :—(1, 2) 54° 44′; (3, 4) 90°. (5, 8, 9, 12) 73° 13′; (6, 7, 10, 11) 30°.

Section 111.

a :—(11) 19° 6′; (3) 40° 54′; (9, 10) 60°; (8) 79° 6′; (7) — 19° 6′; (4) — 40° 54′; (5, 6) — 60°; (12) — 79° 6′.

b :—(1, 5, 9) 90°; (3, 4, 7, 8, 11, 12) 61° 53′; (2, 6, 10) 19° 25′.

Section 212.

a :—(11) 23° 12′; (3) 36° 52′; (9) 60° 57′; (6, 8) 71° 34′; (7) — 23° 12′; (4) — 36° 52′, (5) — 60° 57′; (10, 12) — 71° 34′.

b :—(5, 9) 82° 11′; (1, 3, 4) 74° 13′; (8, 12) 65° 54′; (7, 11) 47° 7′, (2) 35° 15′; (6, 10) 17° 43′.

2. Fig. 7 (Plate 4) is from a photograph (magnification 60 diams.)* of a section of the Coahuila meteorite.

It is an example of a section corresponding with PVUT of fig. 5, the traces parallel to 1, 2 being placed vertical. It contains traces parallel to all the directions of that figure, with the exception of 3, 4, which are absent on this element of the surface. In addition, there are traces parallel to the cube edges, *e.g.*, PT and TU. There arise from sections of rods or plates of rhabdite which are abundant in Coahuila.

Similar low power photographs of sections parallel to other planes were taken but are omitted from consideration of space. Examples of traces on other planes of section occur incidentally in subsequent figures.

The phenomena observed when bands cross one another are described and discussed later.

3. The most direct evidence that the material within the bands is in twin orientation with respect to the matrix is obtained from an examination of the figures produced simultaneously on both by an appropriate etchant.

Of the etchants tried, we found that the most suitable for our purpose was a dilute solution of copper ammonium chloride, the best concentration and time of etching being found by trial in each particular case.

With this etchant the predominant etching planes were parallel to {110};

---

* The magnifications given in the text refer to the original photographs. In the accompanying Plates these are reduced to four-fifths of the previous magnification.

whereas, with dilute nitric acid, the predominant etching planes are parallel to {100}.

This difference between the two etchants is seen very clearly from figs. 8 and 9 (Plate 5), which are photomicrographs (magnification 1500 diams.) of parts of the same {110} section of the meteorite, etched by nitric acid and by copper ammonium chloride, respectively.

In the former the etchant has produced long ridges and troughs with {100} planes as their sides; while, in the latter, flat-bottomed pits with rhombic dodecahedral bounding planes have been developed.

The effects shown in these figures are the result of relatively deep etching. In most cases, however, relatively light etching serves best to show the respective orientations of matrix and bands. By suitable variations of the concentration and time of action of the etchant, the development of individual pits can be followed in detail and very symmetrical etch figures (negative crystals) are obtained.

From observations of this kind it was found that, with copper ammonium chloride as etchant, the incipient pits frequently show facets of forms other than {110}. Of these, the commonest belong to the form {112}. Indeed, facets of this type are often as strongly developed as those of {110} in incipient pits. As the pits develop, however, the {110} facets usually become predominant.

Facets belonging to the forms {100} and {111} are seldom as strongly developed as those belonging to the forms {110} and {112}; but they are rarely absent from every pit upon a particular section.

Illustrations of these remarks appear in subsequent figures.

4. The simplest way of testing whether the twin relationship exists between the orientation in a band and that in the neighbouring matrix is to make a section of the material perpendicular to the {112} plane to which the band is parallel and to compare the pits produced simultaneously on band and matrix by the same etchant. In such case, as can be seen at once by the aid of a figure, *e.g.*, fig. 4 (Part I), the etch-pits in the band and those in the matrix, if similarly developed, should be mirror images of one another in planes represented by the traces of the bands.

Of possible planes of section of this type, the simplest are of the forms {111} and {110}.

The geometric relations to be expected between the pits in the band and those in the matrix, in these two cases, can be derived in the way indicated in figs. 10 and 11.

5. In the auxiliary diagram at the top of fig. 10, the lines PVU and UY indicate a projection, on a {110} plane, of one-half of the lattice cube of fig. 6.

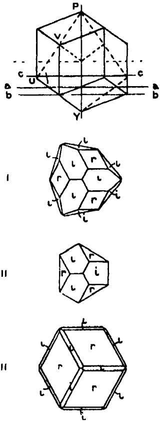


FIG. 10.

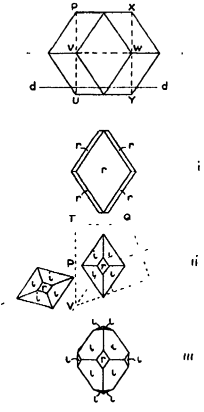


FIG. 11.

The vertical line PY represents a {112} plane perpendicular to the plane of projection. The dotted half-cube to the right of PY indicates the orientation of the cube lattice in the band, assuming it to be in twin relation with the matrix as in fig. 4 (Part I).

The continuous lines in the diagram indicate the orientation of a rhombic dodecahedron formed by the {110} planes of the matrix. This dodecahedron in combination with an icositetrahedron (deleted from the diagram to avoid confusion) was used to obtain the diagrams immediately below it. These are projections upon the {111} plane perpendicular to PY, the plane of the band, of the surfaces obtained by sectioning the combination of dodecahedron and icositetrahedron by octahedral planes represented by the lines *aa* and *bb* respectively.

Fig. 10, i thus represents the theoretical form of a moderately deep pit and fig. 10, ii that of a shallower pit upon a {111} section of the matrix, viewed from above.

Fig. 10, iii represents the form which a relatively deep pit (*e.g.*, with section plane cc) will show, if the development of the {112} facets is insignificant compared with that of the {110} facets.

6. The extent to which these theoretical diagrams accord with the facts can be seen by comparing them with the photomicrographs of fig. 12.

Owing to the high magnifications (*e.g.*, 1800 diams.) required, it was impossible to focus, simultaneously, the contour of a pit and the facets which it exhibited. Consequently, a single photograph of a particular portion of the surface does not convey all the information, about the geometry of the pits upon it, which visual inspection, with continuous adjustment of focus, affords. It was usually possible, however, to find, by trial, approximations to focussing sufficiently close for our present purpose.

In Plate 4, fig. 12A, a Neumann band, corresponding with the plane 1 of fig. 6, runs down the centre of the photograph. Near the middle of the band is a pit of which the form is intermediate between those of the theoretical pits i and ii of fig. 10. In the matrix on either side of the band will be seen pits of similar character, but smaller and, therefore, shallower and approximating more closely to ii of fig. 10 than the pit near the middle of the band.

The mirror image relation between the latter pit and those in the matrix is very clearly seen.

The large pits at the top of the photograph, one wholly in the matrix and the other partly in the matrix and partly in the band, are so deep that the internal facets are out of focus. Their contours are not very regular, but contour angles corresponding with those of fig. 10, i are easily recognisable.

The right-hand boundary between the band and the matrix also shows the mirror image relation by the symmetrical etching on the two sides of the bounding plane. In some cases a small pit develops equally on the two sides of the bounding plane and the mirror image relation between the two halves is very clear.

The peculiarities of the lattice at the boundary, already pointed out in § 9, Part I, probably increase the ease with which etching begins at the junction between a band and the matrix.

A noteworthy fact about bands of the type shown in this photograph is that, apart from the pits, the general appearance of the material of the band is identical with that of the matrix. In consequence, the boundaries look, at first glance, as if they represented two narrow furrows passing across the matrix. This effect is due to the fact that the invisible facets, formed upon the band and matrix at the beginning of the etching process, make the same angles with the surface in the two cases and are therefore equally reflecting. Examples of cases in which this condition does not hold, and of the consequences, will be seen in later photographs.

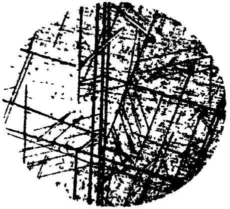
The reason why the {112} facets, in the photomicrographs of fig. 12, show up brightly compared with those of {110} can be seen from the diagram at the top of fig. 10. This shows that while the {110} facets in question make angles of about 35° with the surface, the {112} facets meet it at about 19½°.

7. Fig. 12B is included as an example of a case in which a relatively broad band like that of fig. 12A, is crossed by another. In this case the crossing band is also perpendicular to the plane of the paper and corresponds with the plane 9 of fig. 6. The top right-hand portion of this section is disturbed by the presence of a rhabdite crystal. The "deflexion" of the thin band as it passes through the broad one into a plane corresponding with plane 5 of fig. 6 is discussed later. Owing to its thinness we have not succeeded in obtaining clear etch-pits upon this band.

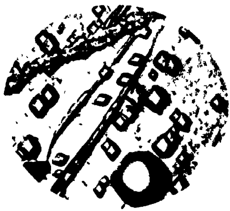
Figs. 12C and 12D are given to show a case of a deep pit corresponding with fig. 10, iii. In the former the hexagon with dark interior is obtained by focussing upon the surface, and in the latter the presence of narrow {112} facets is revealed by deeper focussing. The interior of a large pit is frequently obscured by copper, deposited from the etchant. This can be dissolved by an aerated solution of ammonia.

8. Fig. 13 (Plate 4) is a photomicrograph giving an example of the mirror image relation, when the plane of the section is a {110} plane of the matrix, perpendicular to the plane of the band.

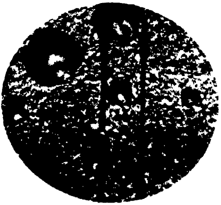
The plane of section and the orientation of the photograph can be described most conveniently by means of figs. 5 and 6. The {110} plane of the section corresponds with the plane TVWQ of fig. 6 and the photograph is oriented so that the direction TV, in which the perpendicular cube face cuts the plane of the photograph, is vertical. The Neumann band, which is perpendicular to the plane of section, corresponds with the plane 3 of figs. 5 and 6, its trace being inclined at approximately 35° to the vertical. It is a relatively poor example of bands of this type, a broader and much more regular one appears, in another connection, in a later photograph (fig. 22A, Plate 9); but it happens to contain good examples of pits showing the mirror image relation with



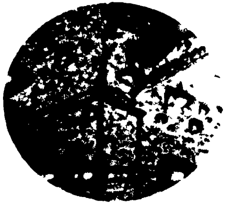
(100) Fig. 7 (× 48)



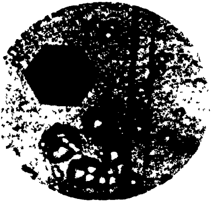
(110) Fig. 13. (× 1680)



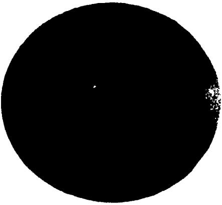
(111) Fig. 12 a (× 1360)



(111) Fig. 12 b (× 1680)



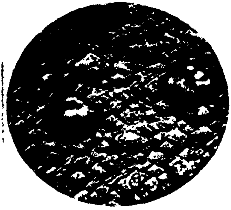
(111) Fig. 12 c. (× 1360)



(111) Fig. 12 d. (× 1360)

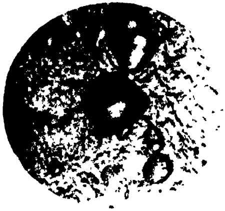
(Facing p. 492.)

(110) Fig. 8. (× 1200)
Etchant Nitric Acid.

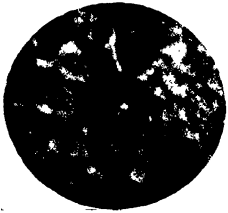


(110) Fig. 9. (× 1200)
Etchant Copper Ammonium Chloride

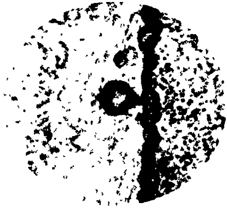
(221) Fig. 16, i (× 1680)



(221) Fig. 16, ii A (× 1680)

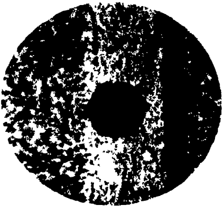


(221) Fig. 16, ii B. (× 1680)

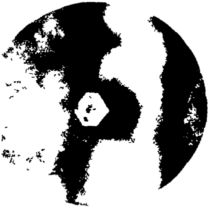


(100) Fig. 16, iii A (× 1680)

(100) Fig. 16, iii b. (× 1680)



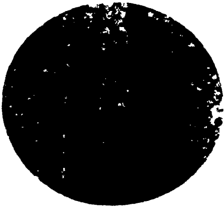
(100) Fig. 16, iv a. (× 1680)



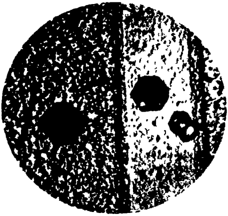
(100) Fig. 16, iv b. (× 1680)

(221) Fig. 16, v. (× 1680)

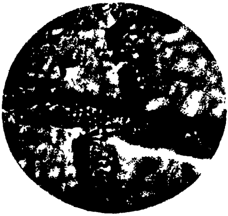
(221) Fig. 16, vi. (× 1680)



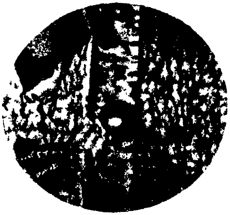
(100) Fig. 16, vii. (× 1680)



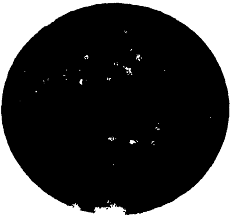
(100) Fig. 18 A (× 1680)



(221) Fig. 18 B (× 1440)



(221) Fig. 18 C. (× 1680)



(221) Fig. 18 D (× 1680)

respect to others in the matrix. These pits and their shapes can be identified more easily by reference to the theoretical diagrams of fig. 11.

These diagrams were derived in the same way as those of fig. 10, the auxiliary diagram being at the top, with the different possible forms for the pits below it.

Fig. 11, i, to be compared with fig. 9, represents the case of a pit which shows rhombohedral facets only. Figs. 11, ii and iii give the forms of pits in which {112} facets are more strongly developed than those of {110}, ii representing a shallower pit than iii. To the left of the former is shown the theoretical orientation, with respect to this pit, of a pit in a band parallel to the plane 3, assuming the band to be in twin orientation with respect to the matrix. The correspondence between these two theoretical pits and those of fig. 13 is unmistakable.

It is of interest to note that the orientation of fig. 13, described above, can be checked by reference to another Neumann band running across the photograph at 65° to the vertical and at 30° to the band just discussed (*cf.* data for traces of {112} planes upon a {110} plane given on p. 487). This band is not perpendicular to the plane of section and, therefore, does not etch in the same manner as the matrix.

9. In the case of a {110} section of the matrix, two bands, and in the case of a {111}, three bands, are perpendicular to the surface. When, however, the plane of section is a {100} plane of the matrix, no band meets the surface perpendicularly.

Consequently, in this case, the simple mirror image relation between the pits in a band and the pits in the matrix is never exhibited. The relationships to be expected now are indicated in fig. 14.

The upper left-hand diagram (*a*) shows two half-cubes in twin position, the plane of projection being a {110} plane of each of them. PGY is the trace of the {112} plane, common to both, perpendicular to the plane of the paper.

The half-cube PVUY may be taken to represent the orientation of the lattice in the matrix. The half-cube PHEY will then represent the orientation of the lattice in a band.

The upper right-hand diagram (*b*) shows the two half-cubes projected upon the {100} plane of the matrix PVUT. The plane perpendicular to the paper represented by YCX in (*a*), parallel to the plane of PVUT, can be seen to be a {221} plane of the band lattice. Consequently, on a section of the matrix parallel to PVUT, the band "outcrop" will present a {221} plane of its own lattice. The direction of the outcrop will be parallel to the diagonal TV of the cube face PVUT of the matrix lattice.

This case is represented in the diagram (a') at the bottom of fig. 14, the directions of the projections of the cubic axes in the band NN being shown by

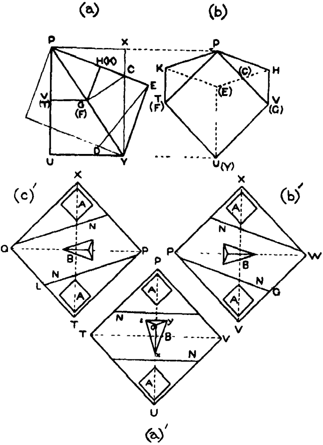


FIG. 14.

the lines Ox, Oy and Oz, of which Ox is perpendicular to TV. Two of the cubic axes in the matrix are parallel, respectively, to the sides of the square A; while the third is perpendicular to the plane of the paper. The projections of cubic etching pits could thus take the forms represented by A and B, respectively.

The band represented by PHEY can, however, outcrop in two other ways upon a cube face of the matrix. It may outcrop upon the cube face represented by PVG and also upon that represented by PTF.

Taking the first case, it can be seen from diagrams (a) and (b) that PGC represents a second {221} plane of the band lattice and that PVG and PGC are co-planar. Consequently, the band again presents a {221} plane when the matrix is sectioned parallel to this {100} plane of its lattice. This case is represented in the diagram (*b′*).

The solution of the corresponding problem for the outcrop of the band upon the remaining cube face of the matrix is shown in the diagram (*c′*).

It will be seen that every Neumann band can be dealt with in the way shown in fig. 14, and that the outcrops of all the bands upon a particular {100} section of the matrix conform to one or other of the types represented by *a′*, *b′* and *c′*.

It will also be seen that, if one had the time and patience necessary to obtain and identify the pits on all the bands upon a single cubic section of the matrix, it would be possible to discover, from them alone, whether or not the bands were parallel to {112} planes of the matrix and twinned in the way already specified.

It will suffice here to give examples of the results obtainable. These are got by comparing the forms of the etching pits observed upon {100} and {221} planes with the theoretical forms deduced by the method employed earlier.

10. It is convenient to begin with the theoretical forms to be expected in pits, showing facets of different kinds, upon a section parallel to a {221} plane of the cubic lattice, the selection having been made after examination of pits obtained experimentally.

The auxiliary diagram at the top of the left-hand column of fig. 15 is a projection of a rhombic dodecahedron oriented so that a {221} plane of the same lattice, represented by the line HK, is perpendicular to the plane of the paper. This dodecahedron was used, in conjunction with other forms deleted to avoid confusion, to obtain the diagrams below it. Diagram i is the theoretical form of a shallow pit showing dodecahedral and octahedral facets. Diagram ii represents a deeper pit in which the predominating {110} facets are accompanied by {100} facets as well as by those of {111}. Diagram iii shows the theoretical form of a still deeper pit exhibiting large {110} facets and small {112} facets; while diagram iv is that of a deep pit exhibiting a combination of {110} and {100}.

The auxiliary diagrams at the top and bottom of the right-hand column represent combinations of the rhombic dodecahedron, oriented in the same way as that at the top of the left-hand column, with the icositetrahedron {112}. In the diagram at the bottom the development of the {112} facets is much more pronounced than in that at the top.

Diagrams v and vi were derived from the upper auxiliary diagram and diagram vii from the one immediately below it. Diagram vi represents a shallower pit than diagram v. Diagram vii represents a relatively deep pit, that of a shallow one of the same type having the contour represented by the dotted straight lines within the diagram.

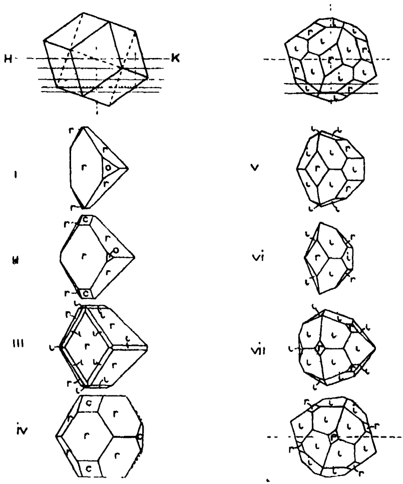


FIG. 15.

From what precedes, pits of any or all of these types are possible, either upon a {221} section of the matrix, or upon a Neumann band outcropping upon a {100} section of the matrix.

11. The photomicrographs reproduced in fig. 16 (Plates 5, 6) are examples of the results obtained. The magnification was between 1800 and 2000 diams. in each case.

The pit near the centre of fig. 16, i corresponds with fig. 15, i, and is given as an example of a pit showing a good {111} facet. Fig. 16, ii A shows the contour of a pit corresponding approximately with 15, ii. The same pit, differently focussed in fig. 16, ii B, is seen to exhibit facets corresponding with those shown in fig. 15, ii.

Fig. 16, iii A and iii B show the contour and facets, respectively, of a pit corresponding with the theoretical pit of fig. 15, iii. It lies in the Neumann band running vertically across the centre of the photograph, the matrix on the right and left being a {100} section of the meteorite. The outcrop of the band is parallel to the diagonal of a cube face of the matrix and the orientation of the pit corresponds with the theoretical orientation shown in fig. 14 (*a*′).

Fig. 16, iv A and iv B show the contour and facets, respectively, of a pit in a band similarly oriented to that of iii A and iii B. The contour and facets correspond with those of fig. 15, iv.

Fig. 16, v shows the facets of a pit upon a {221} section of the matrix corresponding closely with fig. 15, v.

Fig. 16, vi shows a form of pit of frequent occurrence upon a {221} section of the matrix, intermediate between fig. 15, vi and the shallow pit of fig. 15 vii, the {112} facets being predominant.

Fig. 16, vii shows a pit of the same type upon a Neumann band of the same orientation as those of fig. 16, iii and iv, also outcropping upon a {100} section of the matrix.

It will be noticed that the facets which reflect the incident light most effectively are those which, according to the theoretical diagrams, are most nearly parallel to the surface of the section. The deepest {110} plane, making a small angle (19½°) with the surface, reflects strongly; while the two {112} planes, bordering it on the right and making a slightly smaller angle (17¾°) with the surface, reflect a little more effectively. On the other hand, those rhombohedral and other planes which are inclined steeply to the surface are in deep shadow when the incident light is normal to the surface. In one or two cases the lighting was imperfectly adjusted. In others, as in fig. 16, ii B, the relative obscurity of the facets is due to the precipitation upon them of copper from the etchant.

12. In those photographs of fig. 16 which show {100} surfaces, the etch-pits are not very clearly defined. Two common forms which these assume are closely similar to the theoretical diagrams i and ii on the left in fig. 17. The first represents a pit showing {110} and {112} facets only, the latter strongly developed. The second represents a pit showing large {110} facets and also facets of {100} and {112}, the latter being much less strongly developed in this case than in the first. The octagonal contour of a large pit of this

type is shown in the left half of fig. 18 A which is a {100} surface of the matrix. The right half consists mainly of a Neumann band outcrop running

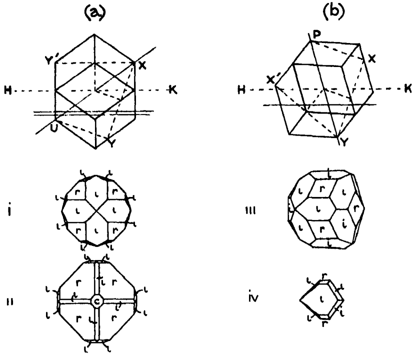


Fig. 17.

across the photograph parallel to a diagonal of the cube face of the matrix. This band contains a large pit of which the contour approximates closely to that of fig. 15, iv. The photograph, as a whole, is thus a particularly clear example of the way in which the etch-pits upon a {100} surface of the material can be used to establish the twin structure of the bands.

13. The remaining photographs of fig. 18 illustrate two other simple relations between the etch-pits upon a band and those upon the matrix.

Fig. 14 has been used to show that the outcrop of a Neumann band upon a {100} surface of the matrix presents a {221} face of the band lattice. It can also be used to show that some of the bands present {100} faces of their own lattices upon a {221} section of the matrix.

A particular case can be seen at once by inspection of the auxiliary diagram (*a*) of fig. 17.

Here the half-cube XYU can be taken to represent the orientation of the matrix lattice. It is in twin orientation with the half-cube XY′U, which will represent the lattice of a Neumann band perpendicular to the plane of the paper, the traces of the band boundaries being parallel to XU. The horizontal line HK represents the trace of a {221} plane of XYU and of a {100} plane of XY′U. The theoretical diagrams i and ii now represent possible forms of the etch-pits in a band of this type—corresponding with the band numbered 2 in fig. 6. The orientations of the cubic axes in bands and matrix will now be those of fig. 14 (*a*′) interchanged.

14. Another form which may be taken by the etch-pits on a Neumann band, outcropping upon a {221} surface of the matrix. can be derived by the method shown in the diagrams on the right in fig. 17.

In the auxiliary diagram (*b*), PXY represents a half-cube of the matrix lattice, HK again representing a horizontal {221} plane of the same lattice. The half-cube PX′Y represents the lattice in the band, the twin plane being PY corresponding with the plane numbered 1 in fig. 6

The diagrams iii and iv give the theoretical forms, obtained as before, of pits showing {110} and {112} facets, iii representing a deep pit and iv a shallow one.

15. The band placed vertically in fig. 18 B represents a case corresponding with fig. 17 (*a*). Cases corresponding with fig. 17 (*b*) are shown in figs. 18 C and D (Plate 7).

In all these cases the orientation of the etch-pits in the matrix is the same , but the etch-pits in the bands of 18 C and D are quite different from those in the band, of parallel outcrop, represented in 18 B.

Fig. 18 C shows shallow pits near the top and a deeper one, out of focus, near the middle. At the centre of fig. 18 D is a deep pit similar to that near the middle of 18 C. The facet at the bottom of this pit, nearly parallel to the surface, is now in focus and is easily identifiable with the corresponding {112} facet of the theoretical diagrams iii and iv of fig. 17 (*b*).

The duplex character of the outcrop of the vertical band in fig. 18 B is discussed later.

16. The utility of the etch-figures produced by copper ammonium chloride, which make it easy to determine the lattice orientations in bands and matrix, is independent of the manner in which they are produced.

It is, nevertheless, interesting to speculate as to why this etchant is so different in its action from a solution of nitric acid.

The atomic density in the {110} planes of the meteorite is greater than that in the {100} planes in the ratio $\sqrt{2} : 1$ and than that in the {112} planes in the ratio $\sqrt{3} : 1$. This might suggest that action would occur most readily upon {110} planes and least readily upon {112} planes.

With nitric acid, in which the product of the interaction is a gas liberated upon the surface of the material, the progress of the etching is subject to the protection which this gas affords to the undissolved portion of the material. This protection would be more effective upon a {110} surface than upon a {100} surface, because upon the latter, as can be seen from a diagram, the removal of one layer of atoms leaves the succeeding layer less effectively obscured by gas than would be the case if the interaction proceeded upon a {110} surface.

It may be that, with copper ammonium chloride as etchant, the deposit being a solid and therefore tending to crystallise, a new factor is introduced.

Copper crystallises in the face-centred cubic system with a lattice constant of 3·60 Å.U. It may be significant that the atomic density in the {110} planes of the meteorite is the same within about 4 per cent. as that in the {111} planes of the copper lattice, and that, similarly, the density in the {112} planes of the one is within about 5 per cent. of that in the {113} planes of the other. If, for example, the attack by the etchant were to proceed in such a way that the copper replaced the matrix metal along the {110} planes of the lattice, the growth of a crystalline deposit would be facilitated.

We have made a preliminary X-ray examination of the deposit and have found distinct evidence in some cases that the copper does, in fact, deposit with a {111} plane of its lattice parallel to the {110} plane of the matrix. The planes along which the etchant unbuilds the matrix lattice would thus appear to be those in which the arrangement of the matrix atoms most resembles an arrangement of the atoms in the copper lattice. These experiments are being continued.

The Mode of Formation of Neumann Bands. Part III.—*The Movement from which the Twinning Results.*

By S. W. J. Smith, F.R.S., A. A. Dee, B.Sc., D.I.C., and J. Young, B.Sc., F.R.A.S., Physical Laboratory, University of Birmingham.

(Received September 24, 1928.)

[Plates 8, 9.]

1. In the preceding Part II we have shown that Neumann bands are lamellæ, parallel to the {112} planes, which are in twin orientation with respect to the matrix.

It remains to show how the nature of the movement, from which a band results, can be determined.

For this purpose we have to find a means of measuring the extent and the direction of the movement of the matrix on one side, with respect to that on the other, when a band of measurable width is formed.

Since all the bands are not formed simultaneously, it must happen that the tracks of some will pass, through the matrix, across the tracks of others formed earlier. The displacement of the parts of an earlier track, with respect to one another, produced during the formation of a later band, can then be used to determine the movement from which the later band results.

2. There is one type of crossing, of one band by another, to which this method does not apply. It is convenient to begin with the consideration of this type, since it throws light upon the others.

The {112} planes of the cubic system can be grouped in four sets of three. The three planes of each set intersect along a trigonal axis and each plane is inclined at 120° to the other two; it is convenient to describe them as "associated" planes.

The four sets can be visualised by the aid of fig. 6.* They are, respectively, the planes 1, 5 and 9; 2, 8, 12; 3, 6, 11; 4, 7, 10. Each plane of the first set is, in fig. 6, perpendicular to the plane of the paper as is the trigonal axis through P in which they intersect. The trigonal axes associated with the other sets are parallel to UX, VQ and WT, respectively.

3. Fig. 19 is intended to represent the problem which arises when a band parallel to one of a set of associated planes is produced in a matrix already crossed by a band parallel to another plane of the same set.

---

* Part II.

The trigonal axis in which the planes intersect is perpendicular to the plane of the paper, which is parallel to a {111} plane of the matrix. The atoms represented by dots are in the plane of the paper. Those represented by circles are in the parallel atom-plane immediately below it; while those represented by crosses are in the next succeeding atom-plane, the sequence being repeated throughout the material.

The band formed first is represented by AAD_1D_1 and may be assumed to be produced by a displacement downwards of the material to the right of the line AA, the graduated slip occurring between AA and D_1D_1 and the full slip existing everywhere to the right of the latter. We now suppose stresses to operate upon the material which tend to produce a band, similar to the first, in the direction BB inclined at 120° to AA, by slip of the material downwards in the region lying between BB and the lower part of the diagram. In order to keep the diagram as simple as possible, we may suppose the extent of the assumed slip to be the same for the second band as for the first.

In the absence of the first band, the second would go straight across the material between the planes BB and D_2D_2 by the propagation across the material of the atomic re-arrangements indicated to the right of the dotted line PQ in the upper right-hand corner of the diagram.

The presence of the first band has, however, the consequence that in the region EFGH the arrangement of the atoms is already that which the forces causing the production of the second band are tending to produce in the same region. The graduated slip, by which the Neumann band lattice is converted into the mirror image of that in the matrix, cannot therefore take place in this region in the way which is possible on either side of it. Unless this part of the first Neumann band, or indeed the whole band, is to take no part in the slipping process, some method by which the downward shear can be transmitted through the band must be looked for. Consideration shows that at least one method exists. This is indicated in the lower part of the diagram.

The second band is assumed to follow the track between B′F′ and D_2′H′ until it reaches the first band. The downward shear is then transmitted through the latter between F′E′ and H′G′ by the graduated movement downwards (in the direction of the trigonal axis) of the atoms in the band lattice. The shear continues through the matrix to the left of the first band, following the track between E′B′ and G′D_2′ shown in the lower left-hand corner of the diagram. It can be seen that the geometry of the lattice lends itself as readily to the production of this sequence of movements as it does to that assumed to occur during the formation of a single uninterrupted band.

The various atoms in the material, below the planes represented by $D_2'H'$, $H'G'$ and $G'D_2'$, have all been displaced downwards, with respect to the corresponding atoms in the material above $B'F'$, $F'E'$ and $E'B'$, by exactly equal amounts. These amounts are the same as if they had been produced during the formation of a straight band equal in breadth to that actually formed.

It is of interest to note that the effect of the two successive "twinning" movements in the region common to the two Neumann bands, represented by $E'F'G'H'$, is to restore the orientation in that region to its original form. This is the result of successive downward shears of the same amount in planes inclined to one another at 60°.

4. The characteristic feature of the movement which fig. 19 represents is that the displacement which accompanies the formation of the crossing band

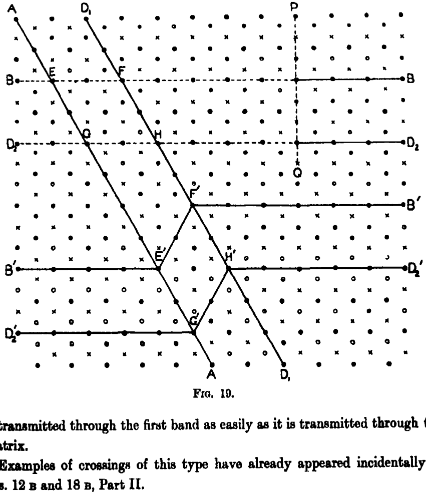


FIG. 19.

is transmitted through the first band as easily as it is transmitted through the matrix.

Examples of crossings of this type have already appeared incidentally in figs. 12 B and 18 B, Part II.

Figs. 20 A and B (Plate 8), which are photographs of bands crossing upon a {221} section of the matrix, are examples of other cases which have been examined.

Fig. 20 A shows the crossing of a broad band, placed vertically in the micrograph, by a narrow one which meets it at an angle of 61°, passes through it at the same angle, and emerges parallel to its original direction. The broad band is known from the etch pits (not in focus in the photograph) to have the same orientation as the bands shown in figs. 17 (*b*) and 18 C and may be taken to correspond with the band numbered 1 in fig. 6 It makes an angle of 74° with the surface of the section, sloping downwards from left to right as in fig. 17 (*b*).

The entering band corresponds with that numbered 9 in fig. 6 (*cf.* data tabulated on pp. 487 and 488). The trace of the track of this band across the broad band has the same direction as the band numbered 5 in fig. 6. It continues parallel to 9 after leaving the broad band. The planes 5 and 9 make equal angles (82° approximately) with the surface of the section. They form, with 1, a set of planes associated with the same trigonal axis. The relation between the crossed and crossing bands is therefore identical with that represented in fig. 19, allowing for the fact that the plane of projection now belongs to the form {221}.

5. Fig. 20 B shows another way in which a crossing of the same type as that just described presents itself upon a {221} surface.

In this case the trigonal axis, with which the crossed and crossing bands are associated, correspond with UX of fig. 6. The crossing band is placed vertically, the direction of its outcrop being parallel to that of the broad band in fig. 20 A. Its orientation is the same as that of the band shown in figs. 17 (*a*) and in 18 B and it corresponds with the band numbered 2 in fig. 6. It makes an angle of 35° with the surface of the section, sloping downwards from right to left as in fig. 17 (*a*).

The crossed band corresponds with the plane numbered 12 in fig. 6, and the trace of the plane in which it is crossed by the entering band corresponds with the plane 8 of the same figure. Each of these planes is inclined at 66° approximately (*cf.* table on pp. 487, 488, Part II) to the surface. The crossing is not symmetrical in this case, the calculated angles of incidence and transmission being $71\frac{1}{2}°$ and 37° respectively. Further, the widths of the two parts of the outcrop are not equal as they are in fig. 20 A. The calculated ratio is $\sin 66° : \sin 35° = \sin 71\frac{1}{2}° : \sin 37° = 1{\cdot}58$. Owing to the smaller angles which this set of planes, particularly 2, make with the surface, the observed angles are more susceptible to errors in the cutting of the section, and in the flatness of the surface, than in the first case; but the correspondence with fig. 19 is again evident.

6. A point of importance in connection with the identification of the outcrops of associated bands upon any section of the matrix is illustrated by fig. 20 B and, more clearly, by fig. 18 B, Part II.

If the Neumann bands are related to the matrix in the way already explained, the orientations of the lattices of every triad of associated bands should be indistinguishable from one another. The angles which the outcrops of these bands make with one another upon the plane of section will vary with the direction in which the matrix is cut. The same will be true of the orientations of the etch-pits with respect to the lines bounding the outcrops of the bands; but the directions of the cubic axes and the orientations of the etch-pits in any one band will always be parallel to those in each of the other two bands. This can be seen to be the case by careful inspection of the crossed and crossing bands (parallel to 12 and 2, respectively) of fig. 18 B.

7. Fig. 20 C is given as an example of the crossing of one associated band by another as exhibited upon a {100} section of the matrix. The crossed band, placed vertically, can be taken to correspond with the plane 2 of fig. 6. The crossing band then corresponds with the plane 8. The track through the crossed band corresponds with the plane 12. The angles of incidence and transmission agree as nearly as they can be measured with the theoretical value, 71° 34′ (*cf.* table, pp. 487, 488).

8. Fig. 20 D contains examples of two peculiarities, not infrequently seen in photographs of the type under discussion. These are of interest because they supplement and confirm our interpretation of figs. 20 A, B, and C.

The orientation of the photograph is the same as that of fig. 20 A and bands of the same type, parallel to 1 and 9, are present in both. The vertical band on the left needs no comment. It shows, relatively imperfectly, a crossing of the same type as that of fig. 20 A. It will be seen, however, that the two bands which encounter the vertical band on the right do not cross it in the normal manner.

9. The lower band apparently stops when it reaches the broader band. Careful examination of the latter shows, however, that its width, below the point at which it is met, is less by a constant amount than its width above that point. It is as if the narrower band has caused a strip to be removed from the broader band and to be added to the matrix, and it is not difficult to see how this effect can have been produced. It may happen, for example, that

the sudden stress, to which the narrow band is due, may be relieved most readily by graduated slip within that part of the material which lies between the narrow band and the lower part of the broad one. This graduated slip forms the narrow band and also reverses the graduated slip in a strip of equal thickness in the broad band, with the result that the original orientation of the material in this strip is restored. At the same time, as will be seen, the whole of the matrix to the right of the bounding planes, represented by the lower side of the narrow band and by the original right-hand boundary of the broad one, has moved parallel to itself through the same distance with respect to the rest of the material.

10. The movements associated with the narrow band, in the upper part of the photograph, can be interpreted in a similar manner. Here the parts of the band on the two sides of the broad one are almost exactly in line and the sudden stress to which they owe their existence has apparently been less local than in the case just considered. It has been relieved most readily by displacement of all the material in the lower part of the photograph with respect to that in the upper. The displacement, greater in this case since the band is wider, has been transmitted by the broad band along two different tracks. It crosses the band almost completely along a track, beginning on the right-hand side, of the same character as that represented in fig. 20 A; but, before it emerges on the left-hand side, it runs along the surface of the broad band, reversing the original graduated slip in the corresponding part of this band and simultaneously reducing its width. At first sight a movement of this kind seems relatively complicated; but it is in reality an example of the comparative ease with which one disturbance can be superposed upon another associated with the same trigonal axis. Even in this case, the part of the broad band in which graduated slip takes place is relatively small. Practically the whole of the band either remains stationary, or moves bodily, along with the matrix in which it is embedded.

11. We come now to the consideration of those cases, most important for our purpose, in which a band, resulting from a particular impulse, crosses a part of the matrix already containing another band not associated with the same trigonal axis.

The new displacement cannot be transmitted across the older band in the way, illustrated in fig. 19, which occasions no change in the alignment of its two parts. What is to be expected in this case can be visualised by the aid of fig. 21.

In this figure, diagram i is intended to represent part of a slab of the material with faces parallel to the plane of the paper. The lines N, N_1 and N′, N_1' are intended to represent the traces of the second Neumann band; produced subsequently to the first, which runs across the material in the direction represented by the lines M, M_1. In order to avoid making the diagrams more complicated than necessary, it is assumed that the {110} plane of the matrix, to which the plane of the second band is perpendicular, is at right angles to the surface of the slab. This {110} plane lies in the plane of the paper in diagram ii, which shows the angle α made by the plane of NN′ with the surface of the slab. As indicated in diagram iii, the band M is assumed to make the angle β with the same surface; its outcrop is assumed, as shown in i, to make the angle θ with the outcrop of NN′. The breadth of the latter is represented by b.

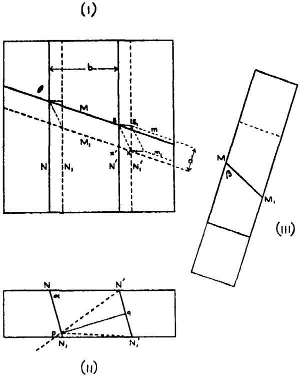


FIG. 21.

After the formation of the band NN′, the surface of the slab to the right of $N'N_1'$, will be no longer co-planar with that to the left of NN_1. Assuming the surface to be so treated that the projecting part is removed, leaving the two parts again co-planar, it is required to find the relation between the traces of the various planes as they present themselves upon the new section.

12. Before this can be accomplished, it is necessary to know how far the surface to the right of N′ moves up or down with respect to that on the left of N when the band represented by $NN_1N'N_1'$ is formed.

Let us assume, provisionally, that the twinning movement is of the type shown in Part I to be that which should occur most easily.

This is indicated by dotted lines in ii, in which the line pN' represents the second {112} plane perpendicular to the plane of the paper, NN_1 representing the first. The line pq is perpendicular to NN_1 and the line pN_1' is equal to pN'. The angle $N'pN_1'$ corresponds with the angle 2ϕ of fig. 1, which, for twinning of the easiest type (*cf.* § 4, p. 482, Part I), is given by $\tan\phi = 1/2\sqrt{2}$.

Hence, with twinning of this type, the surface to the right of N′ is depressed with respect to that on the left of N, when the band is formed, by the amount represented by the vertical distance between N′ and N_1' in diagram ii.

At the same time, the traces of M upon the part of the matrix to the right of $N'N_1'$ move into the positions represented by m and m_1 in i, the former being in the plane represented by N_1N_1' in diagram ii.

When the surface of the slab is reduced to the level assumed by the matrix on the right, the trace of the band M on the portion of the matrix to the left of the band will occupy the position represented by M_1 in i. Consequently, the band first formed will present two straight portions upon the new surface, one to the left of N_1 and the other to the right of N_1' in diagram i, displaced with respect to one another by the amount represented by d in that diagram.

The calculated relation, between this displacement d and the breadth b of the outcrop of the band responsible for it, is

$$d/b = \{\sin^2\alpha\cot\beta + \sin\alpha\cos\alpha\cos\theta\}/\sqrt{2}. \qquad (1)$$

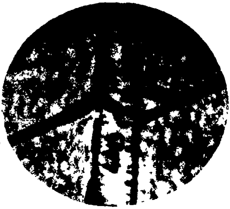
If the band MM_1 had sloped in the opposite direction to that shown, so that M_1 lay on the other side of the normal in iii, the displacement would have been in the opposite direction and given by

$$d/b = \{\sin^2\alpha\cot\beta - \sin\alpha\cos\alpha\cos\theta\}/\sqrt{2}. \qquad (2)$$

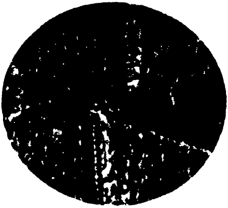
Further, if the movement accompanying the formation of the Neumann band had been upwards instead of downwards on the right (as it would have been if the {112} plane to the left of NN_1 had been parallel to pN_1' instead of to pN') the displacement d, for given values of α, β and θ, would have had the same value as in fig. 21; but its sign would have been reversed.

The method of fig. 21 can, of course, be used to obtain the displacement graphically instead of by calculation.

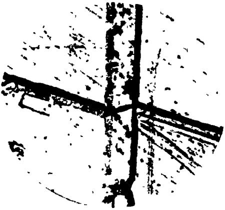
13. The most direct way of applying these results to find the type, to which



(221) Fig. 20 A. (× 1680)

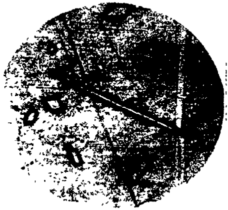


(221) Fig. 20 B. (× 1680)



(100) Fig. 20 c. (× 1680)

(221) Fig. 20 D. (× 1080)

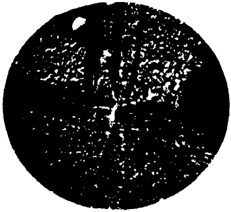


(111) Fig. 23 A. (× 180)

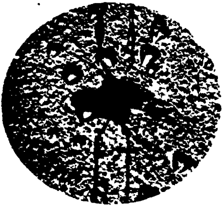
(111) Fig. 23 B. (× 752)

(*Facing p.* 508.)

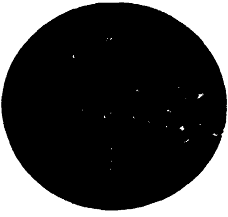
(110) Fig 22 A (1080)



(111) Fig 22 B. (1080)



(111) Fig 22 C (1080)



(221) Fig 22 D (1680)

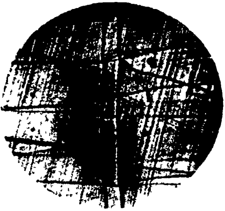


Fig 24 A. (× 1080)

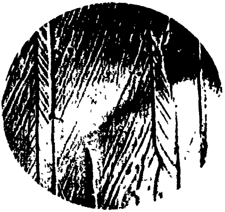


Fig. 24 B. (× 1080)

the movement producing the twinning conforms, is to examine a {110} section of the matrix for a crossing of the required character

In this case, as will be seen from diagram ii, the movement takes place parallel to the plane of section, and the displacement of the trace of a band like M is now parallel to the surface of the specimen It is therefore independent of the angle β, and it gives directly the displacement corresponding with $N'N_1'$ in ii, the breadth of the outcrop of the band producing the displacement being represented by pq, the true width of the band

In this case, therefore, if the twinning is of the most likely type we should have

$$d/b = 1/\sqrt{2},$$

d being upwards or downwards in ii according as the {112} plane, to the left of NN_1 in the matrix, is parallel to pN_1' or to pN'

14 The next simplest case is that in which the band NN′ is perpendicular to the plane of section in i In this case, with $\alpha = \pi/2$ we should find

$$d/b = \cot \beta/\sqrt{2}$$

Examples of these two cases and of the more general one are given in fig 22

15 Fig 22A (Plate 9) is from a {110} section of the matrix The orientation of the photograph can be identified by comparing the pits upon the surface of the matrix with those of fig 13 and with the diagram ii of fig 11 Part II

The broad band sloping downwards from right to left at an angle of 35° to the vertical, belongs to the same family as the similarly oriented band in fig 13, showing the twin etch pits already described For reasons already given, the lustre of the band closely resembles that of the matrix on either side so that only the etch furrows at its boundaries show up distinctly The {110} plane of section corresponding with the plane QTVW of fig 6 the band is parallel to the plane numbered 3 in that figure It corresponds with the band NN′ of fig 21, and the band M which it crosses is the narrow one sloping downwards from left to right at an angle of 35° to the vertical

By comparing the broad band with fig 11 ii, or otherwise, it can be seen that the movement from which it results, if of the easy twinning type, should be an "anti clockwise" one by which the matrix to the right of the band is displaced upwards with reference to that on the left The extent of this movement should be $b/\sqrt{2}$, where b is the breadth of the band Hence, if the broad band breaks the narrow one, it should displace the outcrop of the latter on the right upwards with respect to that on the left by an amount equal to the fraction 0·71 of the breadth of the band's outcrop.

Examination of the photograph will show (*a*) that the gap in the narrow band is exactly equal to the breadth of the broad one and (*b*) that the displacement is of the calculated amount.

It will be noticed also that this type of crossing presents a very different appearance from that observed when one band crosses another associated with the same trigonal axis, the crossed band now acting as an obstruction in the path of the crossing band through the matrix. Further evidence upon this point is given below.

16. Fig. 22B, from a {111} section of the matrix, is an example of the next simplest case of fig. 21. It shows the crossing of a band corresponding with plane 12 of fig. 6 by another corresponding with the plane 1, running vertically across the photograph. The orientation of the latter is such as to facilitate comparison with fig. 21 The cube face of the matrix with whose horizontal diagonal the trace of 1 coincides, slopes downwards from right to left as in fig. 6. The downward slope of the crossed band is towards the lower part of the photograph and the movement which should accompany the formation of the crossing band, if the twinning is of the easiest type, is downwards towards the right. The relation between the bands is thus similar to that represented in fig. 21, except that the band corresponding with NN′ is now perpendicular to the plane of the slab. The displacement of the trace of 12 in the present case should have the same sign as that of M in fig. 21, and, with $\alpha = \pi/2$, the ratio of the displacement of 12 to the breadth of the outcrop of 1 should be

$$d/b = \cot \beta/\sqrt{2} = \cot 62^\circ/\sqrt{2} = 0\cdot 38.$$

Examination of the photograph shows that the observed ratio agrees as nearly as it can be measured with that calculated.

As in the preceding case, none of the other possible types of twinning would give the result shown. The appearance of the crossed band, at the crossing, suggests that it has yielded by slip.

It is very noticeable in fig. 22B and, to a less extent in fig. 22A, that the crossing band " tapers " where the crossed band " obstructs " its path.

17. Fig. 22C shows the same effect very clearly in another way. Here, the " obstruction " is a crystal of rhabdite lying across the track of the movement from which the Neumann band results. The broad band shown has the same orientation as the crossing band in fig. 22B and the obstructing crystal (not in focus) is at the centre of the figure.

During the formation of the band, the matrix to the right has moved bodily downwards with reference to that on the left. This movement, above and below the tapering in the photograph, is accomplished by means of the graduated slip in the band. But within the region where the tapering occurs the graduated slip decreases continuously, and must therefore, presumably, be accompanied by a direct slip. Indeed, unless the rhabdite crystal either moves or breaks when the matrix moves, slip between them seems inevitable.

18. From many other observed crossings of unassociated bands which we have examined, we select that shown in fig. 22D as an example of the more general case represented in fig. 21.

The band placed vertically corresponds with the plane 1 of fig. 20A and the orientation is the same as in that figure, the plane of section being again a {221} face of the matrix-lattice. The other band, inclined at 71½° to the first, corresponds with the plane 12 of fig. 20B. It will be seen that, although the vertical band narrows down and becomes indistinct at the crossing, the parts on the two sides of the crossing are in line. The inclined band is not of as uniform width as it might be for our purpose; but the straight parts of it are parallel to one another although they are not in line. The vertical band slopes downwards from left to right as in fig. 17(*b*) at an angle of 74° (*cf.* table, pp. 487, 488, Part II) to the plane of the section and corresponds with NN′ in fig. 21. The inclined band slopes downwards towards the bottom of the photograph at an angle of 66° and corresponds with the plane M in fig. 21. Taking $\alpha = 74°$, $\beta = 66°$, and $\theta = 71\frac{1}{2}°$ and substituting in the equation (1) we should have in this case

$$d/b = 0{\cdot}35,$$

the direction of d being as shown in fig. 21.

Examination of the photograph shows that the ratio of the displacement of the inclined band to the breadth of the vertical band agrees as closely as it can be measured with the calculated ratio. Thus, again, the easiest type of twinning is shown to be the operating cause of the effects obtained. That shown in the table on p. 482, § 4, of Part I, as the next easiest type, would require a displacement twice as large and in the opposite direction.

Incidentally, by comparing fig. 22D with fig. 20B, the difference between the crossing of unassociated bands and that of associated bands can be seen with particular clearness.

19. Early in the study of meteoric iron, as mentioned in Part I, it was observed that a rhabdite crystal can sometimes be seen to be fractured where

a band crosses it; but, to the best of our knowledge, no attempt has been made to examine such fractures quantitatively. It seemed to us that examination of the relative positions of the parts of selected crystals, fractured as described, might afford further evidence of the character of the movement from which the bands result.

Fig. 23A, Plate 8, is a relatively low-power photograph (magnification 225 diams.) of an area upon a {111} section of the matrix. Owing to the lightness of the etching, the traces of the Neumann bands are not very strongly marked. The photograph is oriented so that the traces of planes parallel to 1 and 2 of fig. 6 are vertical. These traces are parallel to the diagonal of the cube face of the matrix which slopes downwards from right to left, as in fig. 6. Traces of bands parallel to 2 are seen on the right in the photograph. These bands meet the surface at 19½° and are easily distinguished from the matrix by their lustre. On the left of the photograph is the scarcely visible trace of a band parallel to 1, seen more clearly in the photograph 23B. This band, perpendicular to the surface, is difficult to detect owing, as before, to the similarity of its lustre to that of the matrix.

The trace of a band parallel to the plane 7 runs downwards across the section, from left to right, at 19° to the vertical. This band meets the surface at an angle of 62°, the slope of the band being downwards from left to right.

Across the centre of the photograph is the trace of a slightly distorted lamina of rhabdite. The trace makes an angle approximating to 60° with the vertical. The lamina is parallel to the cube face PTQX (fig. 6) of the matrix and meets the surface at an angle of 55°, the slope being downwards towards the upper part of the photograph. Sections of other lath-shaped crystals and rods, also of rhabdite, are seen in various parts of the section.

Fig. 23B, oriented in the same way as that just described, is from a higher-power photograph (magnification 940 diams.), showing fractures and displacements in the left-hand part of the rhabdite lamina, where it is crossed by the bands parallel to 1 and 7 respectively.

The effect of the former can be seen to be a clearly marked displacement downwards of the part of the lamina to the right of the fracture with respect to that on the left. The effect of the band parallel to 7 is to produce a displacement of the opposite sign, *i.e.*, upwards of the part of the rhabdite on the right with respect to that on the left.

20. The significance of the first fracture is easily found. The problem which it presents differs only from that of fig. 22B in the respect that the rhabdite lamina, corresponding with the crossed band 12 of that figure, slopes in the

opposite direction with respect to the normal to the plane of section. In consequence, the displacement should have, as it does, the opposite sign to that shown by 12 in fig. 22B and should be given by

$$d/b = \cot \beta/\sqrt{2} = 0{\cdot}5,$$

since $\beta = \cot^{-1} \tfrac{1}{2}\sqrt{2}$.

In order to bring out the contour of the rhabdite as clearly as possible, the etching was light. Several polishing scratches have not been effaced and the edges of the band 1 are not very clearly defined except for some distance above the fracture in the photograph. The mean of several determinations of its width and of the displacement of the rhabdite, made upon an enlargement of the negative, showed the measured ratio to be within 2 per cent of that calculated.

We have thus another and almost direct method of determining the nature of the movement from which the twinning effect results The break in the rhabdite across the planes of movement (probably along one of the directions in which rhabdite is known to cleave readily) is very clearly shown.

21. In the case of the band parallel to 7, it is less easy to compare the observed effect upon the rhabdite with that to be expected in the light of the preceding results. In this case the {110} plane of the matrix which is perpendicular to the plane of the band is not, as it is in fig. 21, perpendicular to the surface. In consequence, the atomic movements parallel to the trigonal axis in the band, during twinning, no longer occur in planes perpendicular to the direction of the outcrop. The effect registered upon the surface can, however, be regarded as the resultant of two movements, one parallel to the direction in which the band cuts the surface and the other perpendicular to this, in the plane of the band. If the actual movement is D, its components in these two directions, in the present case, are found to be 0·93 D and 0·38 D, respectively.

The displacement upwards d' of the traces of the rhabdite due to the former is given by

$$d'/b = 0{\cdot}93 \sin \alpha \sin \theta/\sqrt{2},$$

and the displacement upwards d'' of the same traces due to the latter is given by

$$d''/b = 0{\cdot}38 \{\sin^2 \alpha \cot \beta - \cos \alpha \sin \alpha \cos \theta\}/\sqrt{2}$$

in which

$$\alpha = 62^\circ, \quad \beta = 55^\circ \text{ and } \theta = 41^\circ.$$

The observed displacement upwards, as in the photograph, should be

$$d = d' + d'',$$

so that we should have

$$d/b = 0{\cdot}44.$$

The photographs, considered together, show that the band parallel to 7 tapers to some extent as it approaches the rhabdite from below; but that its width is practically constant for a considerable distance beyond the point where it leaves the field of view in the high-power photograph. The observed value of d/b again agrees, as closely as it can be estimated, with the calculated value.

22. It is interesting to compare the appearance of the fracture by 7 with that of the fracture by 1 already described. The former, unlike the latter, suggests that one part of the rhabdite has been sheared with respect to the other along the direction of the outcrop of the band. This difference, as indicated by the fact that d' is more than six times as large as d'', arises because the movement producing the fracture is now approximately parallel to the surface and to the direction of the outcrop of the band, instead of being perpendicular to the surface as in the case of the fracture first considered.

The character of the displacement thus confirms the interpretation of its amount.

23. It has been cited against the view that Neumann bands are twins with respect to the matrix, that, when "slip bands" are produced in a region containing a band, the way in which the slip passes through the band affords no evidence of its twin structure.*

We have made a few preliminary experiments upon the Coahuila meteorite in connection with this point, which although it does not bear directly upon the subject of the present communication is yet of great interest. The result of one of these experiments, which are being continued, is shown in fig. 24A, Plate 9.

Here, as in the experiments referred to, the slip passes irregularly across a vertical Neumann band.

In other cases, however, of which fig. 24B is an example, the crossing occurs in such a way as to suggest the twin relation very clearly.

---

* Rosenhain and McMinn, 'Roy. Soc. Proc.,' A, vol. 108, p. 235 (1925).

The Wave Pattern of a Doublet in a Stream.

By T. H. HAVELOCK, F.R.S.

(Received September 18, 1928.)

1. The following paper is a study of the surface waves caused by a doublet in a uniform stream, and in particular the variation in the pattern with the velocity of the stream or the depth of the doublet. In most recent work on this subject attention has been directed more to the wave resistance, which can be evaluated with less difficulty than is involved in a detailed study of the waves; in fact, it would seem that it is not necessary for that purpose to know the surface elevation completely, but only certain significant terms at large distances from the disturbance. Recent experimental work has shown considerable agreement between theoretical expressions for wave resistance and results for ship models of simple form, and attempts have been made at a similar comparison for the surface elevation in the neighbourhood of the ship. In the latter respect it may be necessary to examine expressions for the surface elevation with more care, as they are not quite determinate; any suitable free disturbance may be superposed upon the forced waves. For instance, it is well known that in a frictionless liquid a possible solution is one which gives waves in advance as well as in the rear of the ship, and the practical solution is obtained by superposing free waves which annul those in advance, or by some equivalent artifice. This process is simple and definite for an ideal point disturbance, but for a body of finite size or a distributed disturbance the complete surface elevation in the neighbourhood of the body requires more careful specification as regards the local part due to each element. It had been intended to consider some expressions specially from this point of view, but as the matter stands at present it would entail a very great amount of numerical calculation, and the present paper is limited to a much simpler problem although also involving considerable computation.

A horizontal doublet of given moment is at a depth f below the surface of a stream of velocity c; the surface effect may be described as a local disturbance symmetrical fore and aft of the doublet together with waves to the rear. Two points are made in the following work. One is the variation of the local disturbance with the depth of the doublet, or rather with its relation to the velocity. Roughly, it may be said that the local surface effect changes from a depression to an elevation at a certain speed, which we might have anticipated, to be somewhere about the speed $\sqrt{(gf)}$. A line doublet is first examined, and the surface elevation immediately over the doublet is calculated; it is found to be zero at approximately a speed $0 \cdot 86 \sqrt{(gf)}$. To illustrate the difference for speeds greater or less than this value, curves are shown in fig. 1 for the complete surface elevation when gf/c^2 has the values 4 and $0 \cdot 5$. A three-dimensional doublet is then considered and a similar calculation for the surface elevation immediately over the doublet gives a critical speed of about $0 \cdot 84 \sqrt{(gf)}$.

The second point is the variation of the wave pattern. We may compare it with the pattern due to an ideal point disturbance of the surface of the stream. In that case the approximate evaluation of the integrals by the method of stationary phase gives the system of transverse and diverging waves established in the rear. But in our case there is a variable amplitude factor for the constituent harmonic terms of the integrals, and we notice that the velocity $\sqrt{(gf)}$ has here also a special significance; for the amplitude factor itself possesses an additional stationary value, a maximum, when the velocity exceeds $\sqrt{(gf)}$. The difference this makes in the wave pattern is examined; roughly, at lower speeds the pattern consists chiefly of transverse waves, while at higher speeds the diverging waves become of increasing relative importance. A direct numerical study has been made of the integral for this part of the surface elevation for two values of gf/c^2, namely, 4 and $0 \cdot 5$; graphs are given in figs. 3 and 4 for the surface elevation along various radial lines from the origin, including some outside the limits of the ideal wave pattern.

2. Take Ox in the undisturbed surface of the stream, and Oy vertically upwards, and let the velocity of the stream be c in the negative direction of Ox. Let there be a two-dimensional horizontal doublet of moment M at the point $(0, -f)$. The solution of the problem is familiar as the first approximation for the effect of a submerged cylinder of radius a, if we take $\mathrm{M} = ca^2$. We quote here the complete expression for the surface elevation η in the form used in previous calculations*

$$\eta = \frac{2\mathrm{M}f}{c\,(x^2+f^2)} + \frac{2\kappa_0\mathrm{M}}{c}\int_0^\infty \frac{m\cos mf - \kappa_0 \sin mf}{m^2+\kappa_0^2}\, e^{-mx}\, dm, \tag{1}$$

for $x > 0$, and

$$\eta = \frac{2\mathrm{M}f}{c\,(x^2+f^2)} + \frac{2\kappa_0\mathrm{M}}{c}\int_0^\infty \frac{m\cos mf - \kappa_0 \sin mf}{m^2+\kappa_0^2}\, e^{mx}\, dm + (4\pi\kappa_0\mathrm{M}/c)\, e^{-\kappa_0 f}\sin \kappa_0 x, \tag{2}$$

for $x < 0$, with $\kappa_0 = g/c^2$.

---

* 'Roy Soc. Proc.,' A, vol. 115, p. 271 (1927).

The last term in (2) gives the regular waves to the rear, and the remaining terms the local disturbance which is symmetrical before and behind the origin. The integral in (1) is the real part of

$$\int_0^\infty \frac{e^{-(\alpha+i\beta)u}}{u+i}\,du, \tag{3}$$

where $\alpha = \kappa_0 x$, $\beta = \kappa_0 f$. Asymptotic expansions may be obtained for large values of the parameters, or the integral may easily be evaluated directly by numerical methods when α is not small. For small or moderate values of α and β, (3) may be calculated from

$$-(\mathrm{A}+i\mathrm{B})\,e^{-(\beta-i\alpha)},$$

$$\mathrm{A} = \gamma + \log r + \sum_1^\infty \frac{r^n}{n\,.\,n\,!}\cos n\theta,$$

$$\mathrm{B} = \pi - \theta - \sum_1^\infty \frac{r^n}{n\,.\,n\,!}\sin n\theta,$$

$$r = (\alpha^2+\beta^2)^{\frac{1}{2}}, \qquad \tan\theta = \alpha/\beta, \qquad \gamma = 0{\cdot}5772. \tag{4}$$

Consider the surface elevation at the origin ($x = 0$). Since we have

$$\int_0^\infty \frac{u\cos\beta u - \sin\beta u}{1+u^2}\,du = -\,e^{-\beta}\,\mathrm{li}\,(e^\beta), \tag{5}$$

for $\beta > 0$, where li is the logarithmic integral, we have at the origin

$$\eta = \frac{2\mathrm{M}}{cf}\{1 - \beta e^{-\beta}\,\mathrm{li}\,(e^\beta)\}. \tag{6}$$

Using tables of these functions, we find that η is zero when β is approximately $1{\cdot}35$, or when $c = 0{\cdot}86\sqrt{(gf)}$; when c is less, the value of (6) is negative, while at greater speeds it is positive. To illustrate this point, the surface elevation has been calculated from the complete expressions (1) and (2) for two different cases, $\kappa_0 f = 4$ and $\kappa_0 f = 0{\cdot}5$. The graphs are shown in fig. 1, A being for the smaller value of $\kappa_0 f$ and B for the larger. The ordinates are to the same scale assuming M and f constant and c to be the variable; the abscissæ are in wave-lengths, or more strictly the values of $\kappa_0 x$.

3. Consider now the three-dimensional problem. Take Ox and Oy in the surface of the stream, the current being in the negative direction of Ox, and take Oz vertically upwards. For a horizontal doublet of moment M at the

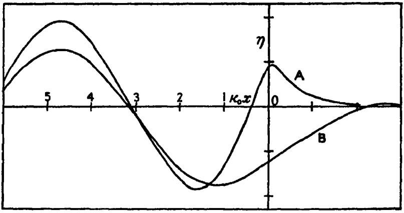


FIG. 1.

point (0, 0, $-f$), we take the velocity potential from a previous paper in the form*

$$\phi = cx - \frac{i\mathrm{M}}{2\pi}\int_{-\pi}^{\pi}\int_0^{\infty} \kappa e^{-\kappa(z+f)+i\kappa\varpi}\cos\theta\,d\theta\,d\kappa$$

$$+\frac{i\mathrm{M}}{2\pi}\int_{-\pi}^{\pi}\int_0^{\infty} \kappa\mathrm{F}(\theta,\kappa)\,e^{-\kappa(f-z)+i\kappa\varpi}\cos\theta\,d\theta\,d\kappa,$$

with

$$\mathrm{F}(\theta,\kappa) = \frac{\kappa+\kappa_0\sec^2\theta+i\mu\sec\theta}{\kappa-\kappa_0\sec^2\theta+i\mu\sec\theta},$$

$$\varpi = x\cos\theta + y\sin\theta. \tag{7}$$

The real part of the expression is to be taken, and further the limiting value as $\mu \to 0$. The surface elevation is obtained from

$$c\frac{\partial\zeta}{\partial x} = \frac{\partial\phi}{\partial z}.$$

After some reduction, ζ is obtained in the form

$$\zeta = \frac{2\mathrm{M}f}{c(x^2+y^2+f^2)^{3/2}} + \frac{\kappa_0\mathrm{M}}{\pi c}\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi}\sec^2\theta\,d\theta\int_0^{\infty}\left\{\frac{e^{i\kappa\varpi}}{\kappa-\kappa_0\sec^2\theta+i\mu\sec\theta}\right.$$

$$\left.+\frac{e^{-i\kappa\varpi}}{\kappa-\kappa_0\sec^2\theta-i\mu\sec\theta}\right\}e^{-\kappa f}\kappa\,d\kappa. \tag{8}$$

Transforming the integral with respect to κ in (8), and taking the limiting value, we obtain

$$2\int_0^{\infty}\frac{\kappa_0\sec^2\theta\cos mf + m\sin mf}{m^2+\kappa_0^2\sec^4\theta}e^{-m\varpi}\,m\,dm, \text{ for } \varpi > 0; \tag{9}$$

---

* 'Roy. Soc. Proc.,' A, vol. 118, p. 28 (1928).

$$2\pi\kappa_0 \sec^2\theta\, e^{-\kappa_0 f \sec^2\theta} \sin(\kappa_0 \varpi \sec^2\theta)$$

$$+ 2\int_0^\infty \frac{\kappa_0 \sec^2\theta \cos mf + m \sin mf}{m^2 + \kappa_0^2 \sec^4\theta} e^{m\varpi}\, m\,dm, \text{ for } \varpi < 0. \quad (10)$$

We have now to integrate with respect to θ, subject to the conditions in (9) and (10). The form of the surface is symmetrical with respect to Ox, so we may write down only the expressions for y positive; and we shall put

$$x = r\cos\theta', \qquad y = r\sin\theta'. \quad (11)$$

We find that the value of ζ can be given by one expression, valid for $0 \leqslant \theta' \leqslant \pi$, namely,

$$\zeta = \frac{2Mf}{c(r^2+f^2)^{3/2}}$$

$$+ \frac{2\kappa_0 M}{c}\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \sec^2\theta\, d\theta \int_0^\infty \frac{\kappa_0 \sec^2\theta \cos mf + m\sin mf}{m^2 + \kappa_0^2\sec^4\theta} e^{-mr|\cos(\theta-\theta')|}\, m\,dm$$

$$+ \frac{4\kappa_0^2 M}{c}\int_{-\frac{1}{2}\pi}^{\theta'-\frac{1}{2}\pi} \sec^4\theta\, e^{-\kappa_0 f\sec^2\theta} \sin\{\kappa_0 r\cos(\theta-\theta')\sec^2\theta\}\, d\theta. \quad (12)$$

This expression is exact, apart from the usual limitation that, at the surface of the stream, we neglect the squares of the additional fluid velocities.

4. The first two terms in (12) represent the local effect which is only of importance in the neighbourhood of the origin. A few preliminary calculations show that, as in the two-dimensional case, it changes from a depression to an elevation about the value $\kappa_0 f = 1$. Considering the elevation at the origin, we have from (12) with $r = 0$,

$$\zeta_0 = \frac{2M}{cf^2} + \frac{4\kappa_0 M}{\pi c f}\int_0^{\frac{1}{2}\pi} \sec^2\theta\, d\theta \int_0^\infty \frac{p\cos u + u\sin u}{p^2+u^2}\, u\, du, \quad (13)$$

where $p = \kappa_0 f \sec^2\theta = \beta\sec^2\theta$.

The integration with respect to u can be expressed in terms of the logarithmic integral, and we obtain finally

$$\zeta_0 = \frac{4M}{\pi c f^2}\left[\frac{\pi}{2} + \int_0^{\frac{1}{2}\pi} p\,\{1 - pe^{-p}\,\mathrm{li}(e^p)\}\, d\theta\right]. \quad (14)$$

The integral in (14) was evaluated approximately for certain values of $\kappa_0 f$ ranging between 1 and 2. The integrand was calculated in each case for a sufficient number of values of p and was then graphed on a base of θ; the value was found by taking values from the graph and using Simpson's rule. In this way it was estimated that ζ_0 is zero at about $\kappa_0 f = 1{\cdot}4$, or $c = 0{\cdot}84\sqrt{(gf)}$. It was also verified that at lower speeds ζ_0 is negative, and at higher speeds positive. For comparison of the maximum local effect with the rest of the surface elevation in two cases discussed later, it may be noted that

$$\begin{aligned} \pi g^{1/2} f^{5/2} \zeta_0 / 4\mathrm{M} &= -1\cdot346, \text{ for } \kappa_0 f = 4 \\ &= +0\cdot860, \text{ for } \kappa_0 f = 0\cdot5. \end{aligned} \tag{15}$$

The local effect at points other than the origin was not calculated, although rough estimates were made for the central line, $\theta' = 0$, from (12) to verify that it falls off in much the same way as in the two-dimensional case; for this purpose the second term in (12) was put into the form

$$(4\kappa_0 \mathrm{M}/\pi c f) \int_0^{\frac{1}{2}\pi} \mathrm{J} \sec^2\theta \, d\theta, \tag{16}$$

where

$$\begin{aligned} \mathrm{J} &= \int_0^\infty \frac{p \cos u + u \sin u}{p^2 + u^2} e^{-\alpha u/\sqrt{(\beta p)}} u \, du \\ &= \frac{p^2}{\rho^2} - p e^{-p} \left[\mathrm{P} \cos\{\alpha\sqrt{(p/\beta)}\} - \mathrm{Q} \sin\{\alpha\sqrt{(p/\beta)}\}\right], \\ \mathrm{P} &= \gamma + \log \rho + \sum_1^\infty \frac{\rho^n}{n \, . \, n!} \cos n\phi, \\ \mathrm{Q} &= \pi - \phi - \sum_1^\infty \frac{\rho^n}{n \, . \, n!} \sin n\phi \, ; \end{aligned} \tag{17}$$

with $\rho^2 = p^2 + p\alpha^2/\beta$, $\tan\phi = \alpha/\sqrt{(\beta p)}$, $\alpha = \kappa_0 r$, $p = \kappa_0 f \sec^2\theta$, $\beta = \kappa_0 f$.

5. Consider now the third term in (12). For computation, we alter the form slightly. We take $\theta' = \pi - \phi$, so that ϕ is the angle the radius vector makes with the negative axis of x, and further we put

$$t' = \cot\phi \, ; \qquad t = \tan\theta. \tag{18}$$

Then this part of the surface elevation, which we may denote by $\zeta - \zeta_1$ is given by

$$\zeta - \zeta_1 = -\frac{4\kappa_0{}^2 \mathrm{M}}{c} e^{-\beta} \int_{-\infty}^{t'} (1 + t^2) e^{-\beta t^2} \sin\left[\alpha (t' - t) \{(1 + t^2)/(1 + t'^2)\}^{\frac{1}{2}}\right] dt. \tag{19}$$

In this form α, or $\kappa_0 r$, is a positive quantity, r being the distance from the origin. The axis of x in front of the disturbance is given by $t' = -\infty$, and (19) is then zero; for the axis of x in the rear of the origin $t' = +\infty$. The usual first approximation to the integral in (19) consists in assuming α large enough so that the only appreciable contributions come from the groups of terms near the positions of stationary phase of the harmonic constituents; this leads to the familiar pattern of transverse and diverging waves contained within radial lines making angles of about 19° 26′ on either side of the negative axis of x, or lines for which $t' = \pm 2\sqrt{2}$. Within the range $\infty > t' > 2\sqrt{2}$, there are two values of t for which the phase is stationary, namely, the roots of

$$2t^2 - t't + 1 = 0, \tag{20}$$

the smaller root corresponding to the transverse wave and the larger to the diverging wave at each point. In the elementary ideal case the constituent harmonic waves have equal amplitudes, but in (19) we have the amplitude factor $(1 + t^2)\, e^{-\beta t^2}$. If $\beta > 1$, this function has a maximum at $t = 0$, and diminishes steadily to zero as t increases. But if $\beta < 1$, there is a minimum at $t = 0$ and a maximum at $t = \{(1 - \beta)/\beta\}^{\frac{1}{2}}$. We may expect then a difference in the wave pattern according as c^2 is greater or less than gf. When $\beta > 1$ and α has moderate values, the main part of (19) comes from small values of t; further, when α is large and the typical wave pattern should be developed, we see that the diverging waves will be relatively small. On the other hand, when $\beta < 1$, there is increasing importance of the diverging waves; and in particular, there will be a value of t', that is a certain radial line, for which the maximum of the amplitude factor coincides with the greater root of (20) for which the phase is stationary. As we are not calculating the wave pattern at large distances we need not put down the general first approximations to (19) by the stationary phase method; we may, however, note the particular cases for $t' = \infty$ and $t' = 2\sqrt{2}$, that is for radial lines along the rearward axis and along the line of cusps of the so-called isophasal lines. For these cases (19) gives, by the usual methods,

$$\zeta - \zeta_l = -\frac{2^{\frac{1}{2}}\pi^{\frac{1}{2}}}{(\kappa_0 r)^{\frac{1}{2}}} \cdot \frac{4\kappa_0^2 \mathrm{M}}{c} \cdot e^{-\kappa_0 f} \cos\left(\kappa_0 r - \frac{\pi}{4}\right),$$

for $t' = \infty$, and

$$\zeta - \zeta_l = -\frac{2^{\frac{1}{3}} 3^{\frac{1}{2}} \pi}{(k_0 r)^{\frac{1}{3}}\, \Gamma(\frac{2}{3})} \cdot \frac{4\kappa_0^2 \mathrm{M}}{c} \cdot e^{-\frac{3}{2}\kappa_0 f} \sin\left(\tfrac{1}{2}\kappa_0 r\sqrt{3}\right), \tag{21}$$

for $t' = 2\sqrt{2}$. We note here the additional factor $e^{-\frac{1}{2}\kappa_0 f}$ in the second case, so far as variation of the amplitude with the depth is concerned; we see that the relative prominence of the so-called cusp waves is only a feature of the limiting case of a point surface disturbance.

6. Returning to the exact integral (19) for this part of the surface elevation, it seemed of interest to make some numerical calculations directly from the integral for points near to the origin, or for moderate values of α. Instead of following the isophasal lines, which are not significant in this region, we have calculated the surface elevation from (19) along certain radial lines. We take in turn the values $t' = \infty$, 3, $2\sqrt{2}$, 2, 1 and zero; these are shown in fig. 2 as A, B, C, D, E and F respectively.

For a given value of t', the value of (19) was found for about a dozen values of α, so that the graph could be obtained with sufficient accuracy for our purpose over a range of α, that is of $\kappa_0 r$, extending from 0 to 18. In each case the value of (19) was obtained by evaluating the integrand at intervals of 0·1 for t for a sufficient range of t until, by reason of the exponential factor, the remaining terms became negligible. Sets of calculations were made for two values of β, that is of $\kappa_0 f$, namely, 4·0 and 0·5; in the latter case it was necessary to take 40 or more values of the integrand in each case, but a smaller number sufficed in the former case. The value of the integral was obtained finally by using Simpson's rule. The collected results are shown in the graphs of figs. 3 and 4, the curves being lettered in agreement with the radial lines of fig. 2.

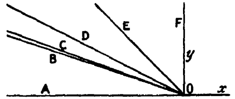


FIG. 2.

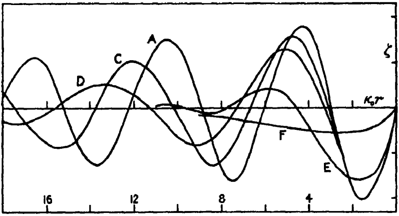


FIG. 3

Fig. 3 is for $\kappa_0 f = 4$, or $c = \frac{1}{2}\sqrt{(gf)}$. Consider first the radial lines within the limits of the ideal wave pattern, namely, A, B, C. In this case, though B and C were calculated separately, there was not sufficient difference to show on the graph without confusion and so B has been omitted; A is the central line and C, at an angle of 19° 26′, would be the cusp line of the simple theory. We may picture the waves in the present case as chiefly transverse waves, slightly curved, and diminishing in height from the central line outwards. The remaining curves are for radial lines outside the usual pattern and show how the wave disturbance is continued in this region. D is for an angle of about 26° 26′ with the rearward central line; it shows an appreciable wave effect, but there are indications that it decreases more rapidly with distance from the origin than for the previous curves. A similar effect, more pronounced, is shown in E and F, for radial lines at 45° and 90° respectively.

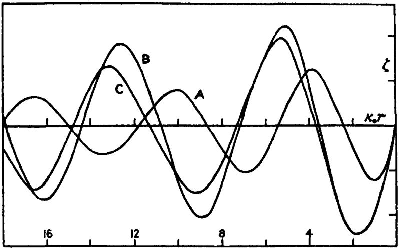


FIG. 4.

Fig. 4 is for $\kappa_0 f = 0\cdot5$, or $c = \sqrt{(2gf)}$. Here, on account of the labour involved in the calculations, only three curves have been drawn; but they bring out the points made in the general discussion. The amplitude along the cusp line C is now greater than along the central line A. Moreover, the greatest amplitude is along the line B, inside the cusp line C, and shows evidence of superimposed diverging waves. The radial line B is given by $t' = 3$. Now for $\kappa_0 f = \beta = \frac{1}{2}$, the maximum of the amplitude factor in (19) occurs at $t = 1$ But from (20), when $t' = 3$ the positions of stationary phase occur at $t = \frac{1}{2}$ and $t = 1$; the latter coincides with the maximum of the amplitude factor and so in this case we should expect a prominent wave along the radial line $t' = 3$, or the line B in the diagram.

A comparison of the curves in figs. 3 and 4 enables us to form some picture of the wave disturbance due to the doublet and the changes that occur as the doublet is brought nearer the surface; in the limit, as far as the wave pattern is concerned, the effect would approximate to the ideal case of a concentrated point disturbance at the surface.

A Symmetrical Treatment of the Wave Equation.

By Prof. A. S. Eddington, F.R.S.

(Received August 30, 1928.)

1. A certain type of matrix has been brought into prominence by P. A. M. Dirac[*] in his theory of the wave mechanics of an electron. C. G. Darwin[†] has commented on the very unsymmetrical form of Dirac's equations, and on the fact that although they are invariant for Lorentz transformations they are not constructed in tensor form. The invariance seems to be of a kind undiscoverable by the usual methods of the tensor calculus and "it is rather disconcerting that apparently something has slipped through the net." In order to throw light on this the following theory of the matrices has been developed; it has several points of interest independently of the application to wave-mechanics. The matrix theory leads to a very simple derivation of the first order wave equation, equivalent to Dirac's but expressed in symmetrical form. It leads also to a wave equation which we can identify as relating to a system containing electrons with opposite spin. This symmetrical method appears to be advantageous in deducing general properties such as the angular momentum of an electron and the conservation of the supposed charge-current vector.

2. *Four-Point Matrices.*—We consider matrices and tensors in four dimensions. The axes are grouped in pairs in three ways denoted by

$$\alpha = 12, 34 \qquad \beta = 13, 42 \qquad \gamma = 14, 23.$$

Thus the points (1, 2) (2, 1) (3, 4) (4, 3) will be called α-points. Besides α-, β- and γ-points there remain (1, 1) (2, 2) (3, 3) (4, 4) which will be called δ-points. We define S_α to be a matrix with 1's at the α-points and 0's elsewhere, so that

$$S_\alpha = \begin{matrix} 0\,1\,0\,0 \\ 1\,0\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,1\,0 \end{matrix} \qquad S_\beta = \begin{matrix} 0\,0\,1\,0 \\ 0\,0\,0\,1 \\ 1\,0\,0\,0 \\ 0\,1\,0\,0 \end{matrix} \qquad S_\gamma = \begin{matrix} 0\,0\,0\,1 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 1\,0\,0\,0 \end{matrix} \qquad S_\delta = \begin{matrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{matrix}$$

We define D_α, D_β, D_γ, D_δ to be diagonal matrices, the leading diagonal being respectively

$$D_\alpha = (1, 1, -1, -1) \quad D_\beta = (1, -1, 1, -1) \quad D_\gamma = (1, -1, -1, 1)$$
$$D_\delta = (1, 1, 1, 1).$$

---

[*] 'Roy. Soc. Proc.,' A, vol. 117, p. 614; vol. 118, p. 351.

[†] *Ibid.*, vol. 118, p. 654.

The elements not on the leading diagonal are zero. Note that

$$S_\delta = D_\delta = 1,$$

that is to say, in a matrix product* the factor S_δ or D_δ makes no difference and can be replaced by 1.

The following results of matrix multiplication are easily verified :—

$$S_\alpha S_\beta = S_\gamma \qquad D_\alpha D_\beta = D_\gamma \tag{1}$$

$$S_\alpha^2 = 1 \qquad D_\alpha^2 = 1 \tag{2}$$

$$S_\alpha D_\beta = -D_\beta S_\alpha \qquad S_\alpha D_\alpha = D_\alpha S_\alpha \tag{3}$$

with similar results obtained by permuting α, β, γ but not δ. The corresponding formulæ for δ are self evident. Denote by (ab) a factor such that

$$\begin{aligned}(ab) &= +1 \quad \text{if} \quad a = \delta \quad \text{or} \quad b = \delta \quad \text{or} \quad a = b\\ &= -1 \text{ otherwise.}\end{aligned}$$

Then (3) is replaced by

$$S_a D_b = (ab)\, D_b S_a \qquad (a,\, b = \alpha,\, \beta,\, \gamma,\, \delta). \tag{4}$$

We have also

$$S_a S_b = S_b S_a \qquad D_a D_b = D_b D_a.$$

3 *Resolution of Tensors.*—The products $S_\alpha D_\alpha$, $S_\alpha D_\beta$, $S_\alpha D_\gamma$ and $S_\alpha D_\delta$ ($= S_\alpha$) are all α-matrices having zero terms except at the α-points, but the first three have -1 at two α-points and $+1$ at the other two. The set of 16 products $S_a D_b$ thus contains 4 α-matrices, 4 β-matrices, etc. The 16 products are independent, *i.e.*, there is no linear identity connecting them. For obviously there could be no linear identity connecting α and β matrices, and the identity if any must be between matrices of the same group, *e.g.*,

$$a_1 S_\alpha D_\alpha + a_2 S_\alpha D_\beta + a_3 S_\alpha D_\gamma + a_4 S_\alpha D_\delta = 0,$$

which is equivalent to

$$S_\alpha\, (a_1 D_\alpha + a_2 D_\beta + a_3 D_\gamma + a_4 D_\delta) = 0,$$

so that there would have to be an identity connecting the four D's. We easily verify that no such identity exists.

It follows that a matrix of 4 rows and columns can be expressed in the form

$$\sum_{ab} t_{ab} S_a D_b. \tag{5}$$

The 16 arbitrary coefficients t_{ab} provide for all possible values of the 16 terms

---

* The product of two matrices a, a' whose elements are $a_{\mu\nu}$, $a'_{\mu\nu}$ is defined by

$$(aa')_{\mu\nu} = \Sigma_\sigma a_{\mu\sigma} a'_{\sigma\nu}.$$

Products are in general non-commutative.

of the matrix, and since the $S_a D_b$ are linearly independent, the equations determining the t's are independent.

We ordinarily describe a tensor by its (1, 2, 3, 4) components $T_{\mu\nu}$, but we can also describe it by its (α, β, γ, δ) components t_{ab}. The t_{ab} are simply sums and differences of four $T_{\mu\nu}$, so that both kinds of components have the same character (*e.g.*, pure numbers, operators). Explicit formulæ determining t_{ab} in terms of $T_{\mu\nu}$ will be obtained in § 14.

4. *Transformation of Tensors.*—A mixed tensor is transformed by change of co-ordinates according to the law*

$$T'^{\nu}_{\mu} = T^{\tau}_{\sigma} \frac{\partial x_\sigma}{\partial x'_\mu} \frac{\partial x'_\nu}{\partial x_\tau}.$$

We write

$$\frac{\partial x_\nu}{\partial x'_\mu} = k_{\mu\nu} \qquad \frac{\partial x'_\nu}{\partial x_\mu} = k'_{\mu\nu}. \tag{6}$$

Then

$$T'^{\nu}_{\mu} = k_{\mu\sigma} T^{\tau}_{\sigma} k'_{\tau\nu}.$$

Since the suffixes on the right are in the proper order for matrix multiplication, we can employ the usual matrix notation which omits them. Thus

$$T' = kTk'. \tag{7}$$

Also by (6)

$$k_{\mu\sigma} k'_{\sigma\nu} = \frac{\partial x_\sigma}{\partial x'_\mu} \frac{\partial x'_\nu}{\partial x_\sigma} = (S_\delta)_{\mu\nu} = (1)_{\mu\nu}$$

or in matrix notation

$$kk' = 1. \tag{8}$$

5. *Rotations.*—The transformation

$$\begin{aligned} x_1 &= x_1' \cos\theta - x_2' \sin\theta \qquad & x_3 &= x_3' \cos\theta - x_4' \sin\theta \\ x_2 &= x_1' \sin\theta + x_2' \cos\theta \qquad & x_4 &= x_3' \sin\theta + x_4' \cos\theta \end{aligned}$$

represents rotations through an angle θ both in the (x_1 x_2) and (x_3 x_4) planes. By (6) we find

$$k = \cos\theta - \sin\theta \,.\, S_\alpha D_\beta. \tag{9}$$

(When a term like $\cos\theta$ has apparently no matrix co-factor, $S_\delta D_\delta$ (=1) can be inserted as a stop-gap. We do not write it explicitly because it is superfluous as soon as k is multiplied by another matrix.) The inverse transformation is a rotation through $-\theta$; hence

---

* We use the summation convention of the tensor calculus for row-and-column suffixes, μ, ν, σ, τ, . . . (occasionally inserting the Σ sign if the summation is likely to be overlooked). But the convention does not apply to matrix-type suffixes a, b, c, d, etc.

$$k' = \cos\theta + \sin\theta \,.\, \mathrm{S}_\alpha \mathrm{D}_\beta .$$

We shall use the term "rotation" with a generalised meaning so as to include any transformation of the form

$$\left.\begin{aligned} k &= \cos\theta + \sin\theta \,.\, i\,(ef)^{\frac{1}{2}}\, \mathrm{S}_e\mathrm{D}_f \\ k' &= \cos\theta - \sin\theta \,.\, i\,(ef)^{\frac{1}{2}}\, \mathrm{S}_e\mathrm{D}_f \end{aligned}\right\}. \tag{10}$$

The reason for inserting the factor $i\,(ef)^{\frac{1}{2}}$ is seen in verifying that k' is inverse to k. We have

$$\begin{aligned} kk' &= \cos^2\theta + \sin^2\theta \,.\, (ef)\, \mathrm{S}_e\mathrm{D}_f\mathrm{S}_e\mathrm{D}_f \\ &= \cos^2\theta + \sin^2\theta \,.\, \mathrm{S}_e\mathrm{S}_e\mathrm{D}_f\mathrm{D}_f && \text{by (4)} \\ &= \cos^2\theta + \sin^2\theta && \text{by (2).} \end{aligned}$$

We have accordingly 16 different "rotations" which correspond to the matrices $\mathrm{S}_e\mathrm{D}_f$ for $(ef) = -1$ and $i\,\mathrm{S}_e\mathrm{D}_f$ for $(ef) = +1$.

6. *Rotation of a Tensor.*—Taking first the case $(ef) = -1$, consider the effect of the rotation $\mathrm{S}_e\mathrm{D}_f$ on a mixed tensor $\mathrm{T}_\mu{}^\nu$. Expressing the tensor in its t_{ab} components we have by (5) and (7)

$$\begin{aligned} \sum_{ab} t'_{ab}\mathrm{S}_a\mathrm{D}_b &= \sum_{ab} t_{ab}\, (\cos\theta + \sin\theta \,.\, \mathrm{S}_e\mathrm{D}_f)\, \mathrm{S}_a\mathrm{D}_b\, (\cos\theta - \sin\theta \,.\, \mathrm{S}_e\mathrm{D}_f) \\ &= \Sigma\, t_{ab}\, \{\cos^2\theta \,.\, \mathrm{S}_a\mathrm{D}_b - \sin^2\theta \,.\, \mathrm{S}_e\mathrm{D}_f\mathrm{S}_a\mathrm{D}_b\mathrm{S}_e\mathrm{D}_f + \sin\theta\cos\theta \\ &\qquad (\mathrm{S}_e\mathrm{D}_f\mathrm{S}_a\mathrm{D}_b - \mathrm{S}_a\mathrm{D}_b\mathrm{S}_e\mathrm{D}_f)\} \\ &= \Sigma\, t_{ab}\, \{\mathrm{S}_a\mathrm{D}_b\, (\cos^2\theta - (af)\,(be)\,(ef)\sin^2\theta) + \mathrm{S}_a\mathrm{S}_e\mathrm{D}_b\mathrm{D}_f\, ((af) - (be)) \\ &\qquad \sin\theta\cos\theta\} \end{aligned} \tag{11}$$

by (4) and (2). (The method of reduction of a long term is to bring all the S's to the beginning, noting each jump of an S over a D, and inserting a bracket of the two suffixes for each such jump.) We have to distinguish two groups of terms—

Group 1.—For these $(af)\,(be) = +1$.

Since $(ef) = -1$, the right side of (11) becomes

$$\Sigma\, t_{ab}\, \{\mathrm{S}_a\mathrm{D}_b\, (\cos^2\theta + \sin^2\theta) + 0\}.$$

Hence these terms are unaltered by the transformation.

Group 2.—For these $(af)(be) = -1$.

The right side of (11) now becomes

$$\Sigma\, t_{ab}\, (\mathrm{S}_a\mathrm{D}_b \cos 2\theta \pm \mathrm{S}_a\mathrm{S}_e\mathrm{D}_b\mathrm{D}_f \sin 2\theta) \tag{12}$$

or since $S_aS_eD_bD_f$ reduces to the form S_cD_d by (1), the result is

$$t_{ab}\,(S_aD_b \cos 2\theta \pm S_cD_d \sin 2\theta).$$

We easily verify that the term in t_{cd} will be

$$t_{cd}\,(S_cD_d \cos 2\theta \mp S_aD_b \sin 2\theta).$$

Hence equating coefficients of the matrices on both sides of (11)

$$\left.\begin{array}{l} t'_{ab} = t_{ab} \cos 2\theta \mp t_{cd} \sin 2\theta \\ t'_{cd} = \pm\, t_{ab} \sin 2\theta + t_{cd} \cos 2\theta \end{array}\right\}, \qquad (13)$$

a transformation which may be regarded as a rotation through an angle 2θ in the plane of these components.

The case $(ef) = +1$, rotation $i\,S_eD_f$, gives the same results, but the circular rotation is now between t_{ab} and $i\,t_{cd}$, equivalent to a hyperbolic rotation or Lorentz transformation of t_{ab} and t_{cd}.

7. *Perpendicular Terms.*—We have seen that a rotation S_eD_f (or $i\,S_eD_f$) leaves a term t_{ab} unaltered if

$$(af)\,(be) = +1, \qquad (14)$$

whereas t_{ab} will be rotated with a conjugated term t_{cd} as in (13) if

$$(af)\,(be) = -1. \qquad (15)$$

Further, c and d are determined by

$$S_c = S_aS_e \qquad D_d = D_bD_f$$

Now

$$(ad) = (ab)\,(af) \qquad (bc) = (ba)\,(be)$$

since a jump of S_a over D_d is equivalent to successive jumps over D_b and D_f. Hence

$$(ad)\,(bc) = (ab)\,(ab)\,(af)\,(be) = (af)\,(be),$$

so that the condition (15) that the two terms are conjugated in rotation becomes

$$(ad)\,(bc) = -1. \qquad (16)$$

Two terms which satisfy this condition will be called *perpendicular*. We also call the two corresponding matrices perpendicular.

The idea of perpendicular and non-perpendicular terms may need some further explanation. Regard the terms t_{ab} as co-ordinates in a 16-space. Two co-ordinate axes define a co-ordinate plane ; *some of the co-ordinate planes in the 16-space are possible planes of rotation and some are not.* We easily deduce from (15) that a given rotation (excluding S_5D_5) rotates 8 of the terms t_{ab} and leaves 8 unaltered. Similarly a given term (excluding t_{55}) is rotated by 8 of the 16 possible rotations S_eD_f and left unaltered by the other 8. Since t_{ab} admits of 8 rotations only, it follows that there must be 7 terms with which it cannot rotate; one of these is always t_{ss}, and the other 6 vary according to the values of a and b.

When we say that the rotation is impossible, we mean that it is impossible without violating the assigned tensor character of t_{ab} which correlates its transformations to transformations of a certain 4-space. It is true that we have only considered "rotations" of this 4-space but it is easily seen that no other transformation could provide the missing rotations in the 16-space. Naturally if we tie a 16-space to a 4-space in such a way that it has only the same number of degrees of freedom, certain motions become impossible Our next procedure is to try to find loci in the 16-space which have more freedom than the rest

8. *Mutually Perpendicular Sets.*—If three terms p, q, r are perpendicular to each other, rotation in the plane pq leaves r unchanged. For let

$$p = S_aD_b \qquad q = S_aS_eD_bD_f \qquad r = S_gD_h$$

so that the rotation from p to q corresponds to S_eD_f. Since r is perpendicular to p and q, (16) gives

$$(ah)(bg) = -1 \qquad\qquad (ah)(eh)(bg)(fg) = -1$$

so that

$$(eh)(fg) = +1,$$

which is the condition (14) that the rotation S_eD_f leaves S_gD_h unchanged

It follows that not more than five terms can form a mutually perpendicular set. For if there were a sixth term it would have to be unchanged by the 10 rotations between the first five terms taken in pairs. But there are only eight rotations which leave a given term unchanged.

It is unexpectedly easy to find a set of five mutually perpendicular terms. Start with any two perpendicular terms and write down the two lists of eight terms perpendicular to each. Three terms are common to the two lists and it turns out that these are perpendicular to one another. From one such set we obtain five others by permuting α, β, γ. The following are the six sets:—

Perpendicular Sets.

S_α	S_α	S_β	S_β	S_γ	S_γ
D_β	D_γ	D_α	D_γ	D_α	D_β
$iS_\gamma D_\beta$	$iS_\beta D_\gamma$	$iS_\gamma D_\alpha$	$iS_\alpha D_\gamma$	$iS_\beta D_\alpha$	$iS_\alpha D_\beta$
$iS_\alpha D_\gamma$	$iS_\alpha D_\beta$	$iS_\beta D_\gamma$	$iS_\beta D_\alpha$	$iS_\gamma D_\beta$	$iS_\gamma D_\alpha$
$S_\gamma D_\gamma$	$S_\beta D_\beta$	$S_\gamma D_\gamma$	$S_\alpha D_\alpha$	$S_\beta D_\beta$	$S_\alpha D_\alpha$

The i's are inserted to show when the rotation between two terms involves the i factor; *e.g.*, in the first set the rotation between D_β and $iS_\gamma D_\beta$ is iS_γ. Another way of putting it is that if the i is transferred to the corresponding $t_{\mu\nu}$ all the terms (so modified) rotate circularly with one another. It will be seen that the matrices with i attached are antisymmetrical, those without i are symmetrical about the leading diagonal. It would, of course, have been permissible to adopt the opposite rule.

The above scheme is exhaustive. Sets of three or four perpendicular terms can only be obtained by breaking up sets of five. Every term (except unity, which is not perpendicular to any term) appears in two columns, and the remaining terms in the two columns are all different; these together give eight terms perpendicular to the term first selected, and we know that there are no more than eight. The terms perpendicular to S_α are either in the first or second column; those perpendicular to D_β are either in the first or sixth column; therefore those perpendicular to both are in the first column. The search for a perpendicular set immediately limits itself to a single column.

A few easily verified properties may be noted. Multiplication of a set by any of the 15 terms partially or wholly destroys its perpendicularity. The five terms + the ten rotations between the terms + unity give the whole set of 16 matrices. (Our convention happens to give an i to the *term* S_aD_b when the *rotation* S_aD_b has none and *vice versa*.) If we multiply a set by one of its terms S_aD_b we obtain the other set containing S_aD_b except that S_aD_b is replaced by unity. This last result is important later on, and I call it the "Coupling Theorem" because it couples two independent sets of four perpendicular terms by a ninth term perpendicular to both.

9. *Invariant Properties.*—Consider a tensor T which has components corresponding to the five terms of the first set, *i.e.*,

$$t_{\alpha\delta},\ t_{\delta\beta},\ it_{\gamma\beta},\ it_{\alpha\gamma},\ t_{\gamma\gamma}, \tag{17}$$

the other 11 components all vanishing. We shall write the 5 components (i's included) as t_1, t_2, t_3, t_4, t_5 and regard them as co-ordinates in a 5-space. By the properties of a mutually perpendicular set rotations are possible in any of the 10 co-ordinate planes of the 5-space, and a rotation in one plane leaves the other three axes unaffected. To this rotation there corresponds a dual "rotation" of the initial 4-space through half the angle.

Roughly we may conclude that if the 4-space is physically isotropic so that its orientation has no absolute meaning, the 5-space also will be indifferent to orientation. Or by substituting it_4 as the fourth co-ordinate, we obtain a world which, apart from its superfluous dimension, is like space-time in that it admits of Lorentz transformations in accordance with the special theory of relativity. We must guard against the impression that the two i's in (17) indicate that the t-world will necessarily have 3 space-like and 2 time-like dimensions. That depends entirely on whether we correlate real or imaginary values of particular t_{ab} with real physical variables when we come to interpret our formulæ physically. The i's in (17) have to do with the linkage between t-space and the initial 4-space, which we shall now call ψ-space. It should be added that "rotation" of the ψ-space has been used in a generalised sense, and "isotropic" might be replaced by a wider term, *e.g.*, "normalised."

Note that the converse is not true. Counting both real and imaginary rotations there are 32 "rotations" of the 4-space, but only 10 of these give internal transformations of the 5-space. The remainder introduce higher dimensions and are merely an artificial complication of phenomena as understood by an observer in 5-space.

The application which we are going to make is to a tensor equation of the form

$$T_\mu{}^\nu\psi_\nu = \psi_\mu, \tag{18}$$

which must be valid for *any* linear transformation of the ψ's. Among these are the 10 rotations above considered. Hence if $T_\mu{}^\nu$ is made up of 5 perpendicular components as above, it can be given any rotation in 5-space and will continue to satisfy equations of the same form.

10. *Square of a Tensor in 5-Space.*—Denote the terms S_aD_b (including i if present) of a perpendicular set by E_1, E_2, .. E_5, and let the tensor be*

$$T = t_1E_1 + t_2E_2 + t_3E_3 + t_4E_4 + t_5E_5. \tag{19}$$

We have

$$E_1{}^2 = 1 \qquad E_1E_2 + E_2E_1 = 0. \tag{20}$$

The second result follows from the condition of perpendicularity (16); the first follows from (2), but attention must be paid to the compensation provided by the i's. Hence if the t's are merely algebraic we have

$$T^2 = t_1{}^2 + t_2{}^2 + t_3{}^2 + t_4{}^2 + t_5{}^2, \tag{21}$$

that is to say, T^2 is a multiple of the unit matrix.† If the t's contain operators which do not commute we must add product terms typified by

$$E_1E_2\,(t_1t_2 - t_2t_1).$$

---

* The t_1's will in some cases differ from the t_{ab}'s by a factor i.

† T^2 denotes TT multiplied matrix-wise. It is not the square of $T_\mu{}^\nu$ in the tensor calculus.

An important case is when t_5 contains no operator, but

$$t_1, t_2, t_3, t_4 = i \operatorname{grad} + V,$$

where V is algebraic. The product terms are then of the form

$$iE_1E_2 (\operatorname{curl} V)_{12}. \tag{22}$$

We see by (21) that a tensor of the form (19) with $\Sigma t_1^2 = 1$ is a square root of the unit matrix, so that (19) is a form likely to be of considerable importance.

11. *The Wave Equation.*—We require a linear equation whose coefficients admit of transformations corresponding to those of the special relativity theory. It is clear that we must adopt

$$T\psi = 0, \tag{23}$$

where T is of the form (19). The only criticism that might be made is that since ordinarily we recognise only four variables (co-ordinates of space and time) admitting of relativistic transformations we are not strictly tied down to T which is a form (the only form) relativistic in five variables. It might be thought that, having chosen E_1. E_2, E_3, E_4 perpendicular, we are free to choose any matrix we like for E_5, since rotations involving it do not interest us. But that is not true; the equation is not invariant if the internal rotations of the first four co-ordinates alter t_5. Hence the fifth matrix must be *non-perpendicular* to the six rotations between E_1, E_2, E_3, E_4; and this is a stringent enough condition to limit it to E_5 or 1—unity being non-perpendicular to everything. Let us choose the latter alternative. Then multiply through by E_5, and by the Coupling Theorem (§ 8) all five terms become perpendicular. Hence the two alternatives correspond to writing the equation in the two equivalent forms $T\psi = 0$, $E_5T\psi = 0$.* The fact that the transformation switches us over to another perpendicular set makes no difference, since we do not specify any particular set to be used in T. It is a curious point that the observed condition that the wave equation must be relativistic in four variables is sufficient to ensure that it is relativistic in five. (But see, however, §17.)

It naturally suggests itself that the second-order wave equation will be

$$T^2\psi = 0, \tag{24}$$

since this is satisfied if (23) is satisfied. By (21) this gives

$$(t_1^2 + t_2^2 + t_3^2 + t_4^2 + t_5^2 + \text{non commutative terms})\ \psi = 0. \tag{25}$$

---

* We have to notice the difference between relativistic *expressions* and relativistic *equations*. Thus T is relativistic in five variables, E_5T in four variables; S_aD_bT is not relativistic at all when the matrix is not one of the five E's. But the *equation* $S_aD_bT = 0$ is relativistic, since it is equivalent to $T = 0$.

Now the orthodox second-order wave equation in simple problems is

$$h^2\left(-\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial t^2}+\frac{m^2c^2}{h^2}\right)\psi=0. \tag{26}$$

Comparing (25) and (26) we make the identification

$$t_1 = ih\partial/\partial x \qquad t_4 = h\partial/\partial t \qquad t_5 = mc.$$

Further, it is generally understood that for more general problems the electromagnetic potential vector V is to be combined with the gradient operator, so that

$$(t_1, t_2, t_3, it_4) = ih\,\text{grad} + (e/c)\text{V} \qquad t_5 = mc. \tag{27}$$

We are now in a position to find the non-commutative terms. Using (22), we obtain

$$\text{T}^2\psi = \{t_1^2 + t_2^2 + t_3^2 - (it_4)^2 + m^2c^2 + (ihe/c)\sum_1^{(4)} \text{E}_\mu\text{E}_\nu(\text{curl V})_{\mu\nu}\}\,\psi = 0 \tag{28}$$

where $(-i\text{E}_4)$ is to be used in place of E_4. The linear equation is

$$\text{T}\psi = \{(ih\,\text{grad} + (e/c)\text{V})\,(\text{E}_1, \text{E}_2, \text{E}_3, -i\text{E}_4) + mc\text{E}_5\}\,\psi = 0 \tag{29}$$

where the scalar product of the bracketed vectors is intended.

12. *Identification of the* E's. There is no restriction on the identification of the E's with particular matrices S_aD_b apart from the condition that they form a perpendicular set. No physical phenomena depend on the choice; physical phenomena depend on the relations of perpendicularity and non-perpendicularity which the matrices are employed to convey. It is no use assigning particular matrices to particular t's; the relativity transformations will make them change partners, interchange being a particular case of rotation. It will often be necessary to substitute particular matrices rather arbitrarily for algebraic convenience in solving special problems; but we hope to show that there is great advantage in treating all the general theory by the symmetrical method of unidentified E's. Meanwhile we make the present digression in order to compare with Dirac's results.

His form of wave equation (*loc. cit.*, p. 615, equation (11)) becomes in our notation

$$(-t_1\text{S}_\gamma\text{D}_\alpha - t_2 i\text{S}_\gamma\text{D}_\gamma - t_3\text{S}_\beta\text{D}_\gamma + t_4 i\text{D}_\alpha + mc\,.\,1)\,\psi = 0.$$

Four coefficients form a perpendicular set, but the fifth member S_β is replaced by unity. Hence (as already explained in § 11) we can multiply through by S_β and the new coefficients will form a perpendicular set—the fourth in our table. The equation now becomes

$$(-t_1\text{S}_\alpha\text{D}_\alpha - t_2 i\text{S}_\alpha\text{D}_\gamma - t_3\text{D}_\gamma + t_4 i\text{S}_\beta\text{D}_\alpha + t_5\text{S}_\beta)\,\psi = 0. \tag{30}$$

Thus Dirac's original equation is equivalent to ours, but in the form $S_\beta T\psi = 0$. The distribution of the real and imaginary matrices in Dirac's identification seems very odd; but there is nothing against its legitimacy, and there may be a practical expediency in employing it in certain problems.* At any rate we have now the whole range of choice before us and can select the matrices most adapted to the problem to be considered. It would seem most natural to avoid introducing more imaginary co-ordinates than necessary by taking the identification

$$(t_1 S_\beta + t_2 D_\gamma + t_3 S_\alpha D_\alpha + (it_4) S_\alpha D_\gamma + (imc) S_\beta D_\alpha)\, \psi = 0. \qquad (31)$$

Since (it_4) corresponds to real time the only imaginary quantities in this equation are (imc) and the imaginary gradient operator in (27). It is of some interest to examine the nature of the "extra terms" in the second-order equation (28)

$$i \sum_{1}^{(4)} E_\mu E_\nu \ (\text{curl } V)_{\mu\nu}.$$

The curl is real, and with our last identification (31) E_1, E_2, E_3, iE_4 are real, so that the whole expression is imaginary. With Dirac's identification the x and z magnetic terms are real and the y terms imaginary, and conversely for the electric terms. The same physical effects must ensue in either case.

It may have been noticed that we have taken $T^2\psi$ for the second-order equation, whereas it has generally been thought necessary to take a form $T'T\psi$ where T' is derived from T by changing the sign of the time-term. The introduction of two different functions is in any case unnecessary, because $T'T$ (*i.e.*, our $T'T$) is a perfect square. It follows from (20) that

$$-TE_4 = E_4 T'.$$

Hence

$$(E_4 T)^2 = E_4 T E_4 T = -E_4 E_4 T' T = -T'T.$$

Actually Dirac starts with a form $T_0 = E_4 T$ in which the matrix of the time-term has been made unity. $T_0' = -TE_4$ gives the same expression with the sign of the time-term reversed. He then uses for his second-order equation $T_0' T_0 = -TE_4 E_4 T = -T^2$. Hence it is the same as ours.

The remainder of this paper is intended to show how the symmetrical method simplifies the derivation of some of the important results already known, and perhaps makes the ideas more comprehensible. The wave equation for two electrons (§ 17) appears at first to break new ground; but I think that its main value is that it introduces in the most natural manner the dissymmetry of electrons which we describe as opposite spin, and gives at once the matrix expression for it. Apart from this the combination is the simple one which would in any case be tried. Its success is contingent on the adequacy of the present mathematical scheme to cover the whole scope of the physical problem.

---

* In semi-statical problems it makes two of the ψ's small compared with the other two, and is therefore advantageous for approximations.

The most symmetrical formulæ are obtained if we put $mc^2 = eV_5$ and write

$$t_\mu = ih\partial/\partial x_\mu + (e/c)\,V_\mu \qquad (\mu = 1, 2, \ldots 5). \tag{32}$$

Here $x_1 \ldots x_5$ are the co-ordinates corresponding to $t_1 \ldots t_5$, so that x_4 is imaginary time and x_5 is a purely formal co-ordinate. This involves inserting an extra operator $ih\partial/\partial x_5$ in t_5; but it makes no difference since none of the physical quantities (ψ, V) are supposed to be functions of x_5. The addition is made for convenience only, and there is no suggestion that the spare co-ordinate shall actually be utilised as a fifth dimension. Its nature is left ambiguous in the one electron problem, but the two electron problem (§ 20) shows that it is not a dimension.

13. *Angular Momentum.*—Quantities M which satisfy MT = TM have a special importance in quantum theory for reasons which we need not here examine. Non-commutative terms arise from two sources: (1) the non-commutative products of the E's; (2) from the differential operators in the t's. In general the quantities M are invented by playing off one source against the other.

We consider the case when there is no electromagnetic field,* *i.e.*, $V = 0$. By (32)

$$(x_1t_2 - x_2t_1)\,t_1 - t_1\,(x_1t_2 - x_2t_1) = -iht_2. \tag{33}$$

The right side appears because t_1x_1 contains $(\partial/\partial x_1)\,x_1$ which is non-commutative; also $x_1\,(t_2t_1 - t_1t_2)$ vanishes because $(\text{curl } V)_{12}$ vanishes. Again by (20)

$$(E_1E_2)\,E_1 - E_1\,(E_1E_2) = -2E_2. \tag{34}$$

Hence multiplying (33) by E_1 and (34) by $\frac{1}{2}iht_1$ and adding, we have

$$M_{12}t_1E_1 - t_1E_1M_{12} = ih\,(-t_2E_1 - t_1E_2)$$

where

$$M_{12} = x_1t_2 - x_2t_1 + \tfrac{1}{2}ihE_1E_2. \tag{35}$$

Operating similarly on t_2E_2 we obtain $ih\,(+t_2E_1 + t_1E_2)$. Operating on t_3E_3, t_4E_4, t_5E_5 everything commutes. Thus

$$M_{12}T - TM_{12} = 0. \tag{36}$$

---

* Most of the practical interest is associated with the fact that the 3 space-components of M preserve the commutative property when there is a central electrostatic field; but this is familiar and need not be repeated here.

The quantities M_{23}, M_{31}, M_{12} are called components of angular momentum. The term $\frac{1}{2}ihE_1E_2$ which is necessary to preserve the commutative property appears unexpectedly and corresponds to the "spin" of the electron.

14. *Determination of the t_{ab}.*—We have postponed the problem of evaluating the components t_{ab} of a tensor T. When T is limited to five perpendicular components the problem is easy; we have in fact

$$E_1T + TE_1 = 2t_1.$$

But we now consider the case of 16 components. Consider first a tensor $T = UV$ which is the product of two vectors. Then

$$t_{ab} = \sum_{\mu\nu} \tfrac{1}{4} US_aD_bV, \tag{37}$$

or inserting suffixes

$$t_{ab} = \tfrac{1}{4} \sum_{\sigma\tau} U_\sigma (S_a)_{\sigma\rho} (D_b)_{\rho\tau} V_\tau. \tag{38}$$

To prove this we must show that $\Sigma\, t_{ab}S_aD_b$ reproduces the tensor UV, *i.e.*,

$$\tfrac{1}{4} \sum_{\sigma\tau} \sum_{ab} \{U_\sigma (S_a)_{\sigma\rho} (D_b)_{\rho\tau} V_\tau\} (S_a)_{\mu\lambda} (D_b)_{\lambda\nu} = U_\mu V_\nu \tag{39}$$

for all values of μ and ν. The function of S_a and D_b on the left is to pair successive suffixes by giving zero factors when the suffixes are improperly paired. D_b may, in addition, reverse the sign. After this pairing we shall be left with an expression of the form

$$\sum_b \tfrac{1}{4} U_\sigma V_\tau (D_b)_{\tau\tau} (D_b)_{\nu\nu}.$$

Now

$$\sum_b (D_b)_{\tau\tau} (D_b)_{\nu\nu}$$

gives two $+1$ terms and two -1 terms if $\tau \neq \nu$, and gives $1 + 1 + 1 + 1$ if $\tau = \nu$. Hence we must take $\tau = \nu$ and cancel the factor $\frac{1}{4}$.

To verify (39) for the component $\mu = 2$, $\nu = 3$, we note that D_b gives zero factor unless $\lambda = \nu = 3$. Then $(S_a)_{23}$ gives zero factor unless $a = \gamma$. Starting again from τ which, we have shown above, $= \nu = 3$, we have from $(D_b)_{\rho\tau}$, $\rho = 3$; $(S_\gamma)_{\sigma 3}$ vanishes unless $\sigma = 2$. The left side is thus

$$\tfrac{1}{4} U_2V_3 \sum_b (D_b)_{33} (D_b)_{33} = U_2V_3 = U_\mu V_\nu.$$

By a similar process all the other components are verified.

The formula (37) also applies when T cannot be split into two factors, the difficulty of putting S_aD_b in the middle of a single letter being merely typographical. When the suffixes are inserted so that position no longer matters we have as in (38)

$$t_{ab} = \tfrac{1}{4} (S_a)_{\sigma\rho} (D_b)_{\rho\tau} T_{\sigma\tau}.$$

15. ***Electric Charge and Current.***—If ψ_μ and ψ^μ are covariant and contravariant vectors we have as in § 4

$$\psi_\mu' = k_{\mu\sigma}\psi_\sigma \qquad \psi'^\mu = k_{\sigma\mu}'\,\psi^\sigma$$

The transformations (10) which we employ are

$$k = \cos\theta + \sin\theta\, \mathrm{S}_e\mathrm{D}_f \quad (ef) = -1,$$
$$k = \cos\theta + \sin\theta i \mathrm{S}_e\mathrm{D}_f \quad (ef) = +1.$$

(1) If $(ef) = -1$, $\mathrm{S}_e\mathrm{D}_f$ is antisymmetrical so that $k_{\sigma\mu}$ is derived from $k_{\mu\sigma}$ by substituting $-\mathrm{S}_e\mathrm{D}_f$ for $\mathrm{S}_e\mathrm{D}_f$. But k' is derived from k by reversing the sign of θ. Hence $k_{\sigma\mu}' = k_{\mu\sigma}$.

(2) If $(ef) = +1$, $\mathrm{S}_e\mathrm{D}_f$ is symmetrical and $k_{\sigma\mu} = k_{\mu\sigma}$. Hence $k_{\sigma\mu}'$ is derived from $k_{\mu\sigma}$ by reversing the sign of the second term.

Both cases are included in the statement that $k_{\sigma\mu}'$ is the conjugate complex to $k_{\mu\sigma}$. It follows that if ψ obeys the covariant law its conjugate complex ψ^* obeys the contravariant law.[†] Hence $\psi\psi^*$ is a mixed tensor.

In general $\psi\psi^*$ will have 16 components j_{ab}. We write

$$\mathrm{J} = j_1\mathrm{E}_1 + j_2\mathrm{E}_2 + j_3\mathrm{E}_3 + j_4\mathrm{E}_4 + j_5\mathrm{E}_5, \tag{40}$$

so that J is the portion of $\psi\psi^*$ which is in our 5-space. The whole tensor $\psi\psi^*$ will be called Υ. Since T and Υ have the same transformation law, being both mixed tensors, the j's have the same transformation laws as the t's for all *internal* transformations of the 5-space. Of course if we were to admit more general transformations this parallelism would fail, since we should then have to introduce the components of Υ which are not included in J.

Accordingly for relativity transformations (j_1, j_2, j_3, ij_4) will transform like $(\partial/\partial x, \partial/\partial y, \partial/\partial z, \partial/\partial t)$ and therefore be a covariant vector in physical space. It is interpreted as an electric charge and current vector, since it is found to be in some respects a counterpart to the charge and current vector of classical theory.

16. *Conservation of Electric Charge.*—By (37) the j's are given by

$$j_1 = \pm\, \psi\mathrm{E}_1\psi^*$$

the lower sign corresponding to antisymmetrical matrices, owing to the i's introduced in j_1 and E_1. The double sign is avoided by writing

$$j_1 = \psi^*\mathrm{E}_1\psi, \tag{41}$$

[†] This is not quite the same as the ψ^* generally employed $(\psi_0{}^*)$ which is related to it by the equation $\mathrm{E}_4\psi^* = \psi_0{}^*\mathrm{E}_4$. The latter refers to real time, and therefore reverses the sign of the hidden i in the co-ordinate x_4. If $\psi_0{}^*$ is used the formulæ for j are less symmetrical.

as may easily be verified. The right side is to be summed, or (equivalently) treated as a scalar product.

The wave equation $T\psi = 0$ gives

$$\left(ih\frac{\partial}{\partial x_1}E_1\psi + ih\frac{\partial}{\partial x_2}E_2\psi + \ldots\right) + \frac{e}{c}(V_1E_1\psi + V_2E_2\psi + \ldots) = 0. \qquad (42)$$

This must be satisfied by the real and imaginary parts separately, so that we may reverse the sign of i throughout. This reversal will turn $E\psi$ into $E'\psi^*$ where $E' = -E$ for antisymmetrical and $E' = +E$ for symmetrical matrices. Both cases give

$$E'\psi^* = \psi^* E.$$

Hence the equation with i reversed is

$$-\left(ih\frac{\partial}{\partial x_1}\psi^*E_1 + ih\frac{\partial}{\partial x_2}\psi^*E_2 + \ldots\right) + \frac{e}{c}(V_1\psi^*E_1 + V_2\psi^*E_2 + \ldots) = 0. \qquad (43)$$

Multiply (42) by initial ψ^* and (43) by final ψ and subtract. We obtain

$$ih\left(\frac{\partial}{\partial x_1}(\psi^*E_1\psi) + \frac{\partial}{\partial x_2}(\psi^*E_2\psi) + \ldots\right) = 0,$$

since $\partial E_1/\partial x_1, \ldots = 0$. By (41) this becomes

$$\operatorname{div} j = 0. \qquad (44)$$

Formally the divergence contains five terms, but it becomes the ordinary divergence of the charge and current vector in the application to the physical world where the ψ's and therefore the j's are functions of the first four variables only.

17. *Wave Equation for Two Electrons.*—There is another mode of approaching the wave equation which has the advantage that it does not perplex us with a spare fifth dimension. Our starting-point will now be, not the relativity of the ordinary physical variables, but the invariance of m, the proper-mass of an electron. Its invariance is not deducible from the theory as previously given; mc ought to be transformable along with the other four t's. We have to impose a condition extraneous to the theory, which forbids these transformations, so that the spare dimension is left unused. Since we have in any case to make this invariance of m a postulate, we may as well start with it.

Accordingly we fix our eyes, as it were, on an oasis in the desert of relativity, and begin with the idea of investigating entities possessing a certain invariant characteristic m. Recognising that matrices must be introduced, we associate

m with an arbitrarily chosen matrix E_5; then any relative physical quantities will have to be associated with other matrices in such a way that their mutual rotations leave E_5 unchanged and so permit m to retain its invariance. This requires that they shall be associated with the eight matrices perpendicular to E_5. Our matrix theory enables us to write these down; they are the remaining members of the two sets in which E_5 appears. Taking E_1, E_2, E_3, E_4 from one set, the others are

$$E_1' = E_5E_1, \qquad E_2' = E_5E_2, \qquad E_3' = E_5E_3, \qquad E_4' = E_5E_4$$

as given by the Coupling Theorem (§ 8). Thus the general form of linear wave equation for which m is invariant when the other coefficients are transformed, is†

$$\mathfrak{X}\psi = \{(t_1E_1 + t_2E_2 + t_3E_3 + t_4E_4) + (t_1'E_1' + t_2'E_2' + t_3'E_3' + t_4'E_4') + 2mcE_5\}\,\psi = 0. \quad (45)$$

We no longer use the observed degree of relativity of the world to discover the wave equation; we have our wave equation, and must use it to deduce the precise character of the relativity of physical variables. We write

$$t_1 = ih\partial/\partial x_1 + (e/c)\,V_1 \qquad t_1' = ih\partial/\partial x_1' + (e/c)\,V_1'. \quad (46)$$

Apart from the fact that we assume that higher derivatives than the first will not be required, this involves no loss of generality.‡ There is no specialisation until we assign meanings to the symbols x_1, V_1, etc., and we have not yet reached that stage. We may note that for transformations of the wave equation V_μ will behave like $\partial/\partial x_\mu$, and V_μ' like $\partial/\partial x_\mu'$, so that V and V′ are covariant vectors for transformations of the unaccented and accented co-ordinates.

If the same rotation or Lorentz transformation is applied to (t_1, t_2, t_3, t_4) and to (t_1', t_2', t_3', t_4') the equation is invariant. For the rotation between E_1 and E_2 is the same as the rotation between E_1' and E_2', *i.e.*, between E_5E_1 and E_5E_2. Or, referring to §§ 6, 7, the transformation of ψ which gives a rotation in the $t_1 t_2$ plane gives also a rotation in three other planes including $t_1' t_2'$. It leaves eight terms unaltered; E_3', E_4' will be among these since they are perpendicular to E_1', E_2'.

---

† The factor 2 in the last term is put in speculatively; I do not know whether it should be there or not.

‡ [*Added October* 13.—I now realise that (45) is the general form of equation for many electrons rather than specially for two. At the time of writing I thought that every co-ordinate would have to be distinguished by a characteristic matrix, so that Dirac's form of equation could apply only to one electron or to two electrons of opposite spin; other cases would involve elaborate extensions. Actually, the case of many electrons is obtained by taking the vectors t, t' to be the sum of a number of gradients (+ an algebraic vector) instead of limiting each to a single gradient.]

Hence we can interpret (x_1, x_2, x_3, x_4) and (x_1', x_2', x_3', x_4') as the co-ordinates of two points *in the same space-time*† and therefore undergoing simultaneously the same relativity transformations. Further we interpret V and V′ as vectors *at* these two points. We call V the electromagnetic vector—which amounts to no more than saying that it is descriptive of something additional to position that we associate with a particular point of space-time.

In general V will be a function of all the eight co-ordinates, and therefore its value will not be completely determined by specifying the point x_μ where it is. Since it depends also on x_μ', we must attach to x_μ' something capable of affecting the electromagnetic vector at another point x_μ. This source of electromagnetic influence which x_μ' is supposed to carry is called an *electron*. Hitherto it has been found adequate to assume that its influence on V is that of a classical electron; but of course the resulting V determines observable phenomena, not by classical electrodynamics but by its occurrence as a coefficient in the wave equation.

Since an electron must also be attached to x_μ, V and V′ will be interchangeable functions, *i.e.*,

$$V = f(x_\mu, x_\mu') \qquad V' = f(x_\mu', x_\mu).$$

Provisionally we shall suppose that $\mathfrak{X}\psi = 0$ is the wave equation for a system containing two electrons.

18. *Dissymmetry of the Two Electrons.*—Multiplying (45) by E_5 we obtain

$$E_5\mathfrak{X}\psi = \{(t_1E_1' + t_2E_2' + t_3E_3' + t_4E_4') + (t_1'E_1 + t_2'E_2 + t_3'E_3 + t_4'E_4) + 2mc\}\,\psi = 0. \quad (47)$$

Denote $E_5 m$ by m'. Then (47) is equivalent to (45) with t_1, t_2, t_3, t_4, m and t_1', t_2', t_3', t_4', m', interchanged. Since ψ satisfies both equations, the group of solutions must involve these two sets of variables symmetrically; and if we associate an invariant electron mass m with the co-ordinates x_μ we must associate m' with the co-ordinates x_μ'.

The meaning is that there is an essential dissimilarity between the two electrons. The difference is exhibited most simply by calculating the angular momenta. By (35)

---

† It will appear later that it is better to regard them as two coincident but not identical space-times. The rotation E_5 transforms one into the other; like the relativity transformations it leaves m invariant.

$$\mathrm{M}_{12} = x_1 t_2 - x_2 t_1 + \tfrac{1}{2} i h \mathrm{E}_1 \mathrm{E}_2 .$$

Similarly

$$\mathrm{M}_{12}' = x_1' t_2' - x_2' t_1' + \tfrac{1}{2} i h \mathrm{E}_1' \mathrm{E}_2' = x_1' t_2' - x_2' t_1' - \tfrac{1}{2} i h \mathrm{E}_1 \mathrm{E}_2 .$$

Thus the electrons at the points x_μ, x_μ' have opposite spins $\pm \tfrac{1}{2} i h \mathrm{E}_1 \mathrm{E}_2$.

It is, however, difficult to grasp what a property defined with reference to T amounts to when $\mathfrak{T}$ is substitued for T. If we repeat the investigation of § 13 with $\mathfrak{T}$ substituted for T, we find that

$$\mathfrak{M}_{12} = x_1 t_2 - x_2 t_1 + x_1' t_2' - x_2' t_1' + \tfrac{1}{2} i h \mathrm{E}_1 \mathrm{E}_2 \tag{48}$$

satisfies

$$\mathfrak{M}_{12} \mathfrak{T} = \mathfrak{T} \mathfrak{M}_{12} . \tag{49}$$

If $\mathfrak{M}_{12}^*$ denotes the same expression with the sign of the last term reversed, $\mathfrak{M}_{12}^*$ commutes with $\mathrm{E}_5 \mathfrak{T}$, which we must regard as an equally fundamental form of the wave equation (*cf.* (47)).

The fact that when the momentum of the two electrons is taken together we still have to add a spin-term, rather spoils the picture of two electrons with opposite spin which ought to cancel out.

The quantities m, m' attached to the two electrons in such a way that they must be interchanged if the co-ordinates are interchanged, may be written

$$m1, \quad m\mathrm{E}_5,$$

where m is a pure number and 1 the matrix unity. The reason for the appearance of these two matrices is that they are the only ones non-perpendicular to the six rotations of space-time, and consequently unaltered by relativity transformations.

This duplexity phenomenon can, however, be appreciated without specifically referring to electrons. The relations of our 16 matrices require that if there is one four-dimensional continuum with respect to which m is invariant there must also be another related to m in a different way. Thus a point is specified by its position in space-time together with a "fourth quantum number" specifying which of the two space-times it is in. We can, so to speak, see the fourth quantum number originating in the scheme of perpendicular sets in § 8, where "worlds" with four perpendicular axes are coupled in pairs to a ninth axis which is invariant for their transformations.

We naturally wonder whether it is essential that the matrices in this theory should be limited to four rows and columns; the number has no immediate connection with the four-dimensions of the world. It is conceivable that the matrices may be extended *ad lib.* according to the complexity of the system under discussion. But when we note how precisely the fourfold matrices give the fundamental characteristics of the world, viz., that it consists of two four-dimensional space-times co-gredient in everything except the phenomenon of duplexity—the fourth quantum number limited to two values—it would seem that these, and only these, matrices play the basal part in it.

19. *General Remarks.*—In the usual nomenclature t_1, t_2, t_3, t_4 are the momenta (including energy) of the electron and m is its proper mass. According to the usual conception of momentum as proportional to mass, the momenta of an electron of proper mass $m' = m\mathrm{E}_5$ should be $t_1\mathrm{E}_5$, $t_2\mathrm{E}_5$, etc., assuming that t_1 has been calculated from field-quantities in the same way as for the first electron. Hence it appears that $\mathfrak{T}$ is obtained from T by taking the momentum of the second electron (distinguished by an accent) and adding it to that of the first.

Thus our provisional wave equation of two electrons is not new in principle. There is no great subtlety in supposing, until the contrary is proved, that an equation with momenta as coefficients is not limited to the case in which the momenta belong to one electron. This, no doubt, is what has generally been assumed or expected. But it is one thing to add the momenta of two electrons in the pious hope that this will serve the purpose ; it is rather different when they first appear in our theory in this combination, and we encounter some difficulty (*i.e.*, a spare dimension) when we try to take them apart.

It is interesting to note the way in which the existence of electrons with opposite spin locks the "fifth dimension," so that it cannot come into play and introduce the absolute into a world of relations. The domain of either electron alone might be rotated in a fifth dimension and we could not observe any difference. But if comparison between the domains is possible, we can notice and forbid the opening out of an angle between them. By forbidding such separation we keep m invariant.

On the Principle of Least Action in Wave Mechanics.

By J. M. WHITTAKER, Edinburgh University.

(Communicated by E. T. Whittaker, F.R.S.—Received August 31, 1928.)

§ 1. *Introduction.*

It is well known that de Broglie's wave mechanics is based on the analogy between the mechanical principle of least action and the optical principle of least time. Schrödinger's original theory† of the hydrogen atom was likewise based on a variational principle, and the method has been used extensively in subsequent developments of the subject. Very recently Darwin‡ has given the form of the principle appropriate to Dirac's theory‖ and has shown that expressions for the current vector and electric density can be deduced from it.

In Dirac's work the electron wave is specified by a four-rowed matrix. The theory has been brilliantly successful in accounting for the "duplexity" phenomena of the atom, but has the defect that the wave equations are unsymmetrical and have not the tensor form. In a recent paper¶ an attempt has been made to surmount this difficulty. Eight wave functions are employed instead of Dirac's four. These are grouped together to form two four-vectors and satisfy tensor equations of the second order. It is shown in § 2 of the present paper that these eight wave equations can be reduced, by addition and subtraction, to the four second order equations satisfied by Dirac's functions, taken twice over; and that, in a sense, the present theory includes Dirac's.**

In addition to the wave equations a complete scheme must include electromagnetic and gravitational equations. These will differ from the equations of Maxwell and Einstein in having "wave" terms instead of "particle" terms for the current vector and material energy tensor.

The object of the present paper is to find these equations and to show that the complete system can be deduced from a "principle of least action"; the wave, electromagnetic, and gravitational equations being found by equating to zero the coefficients of the variations of the wave functions, the electromagnetic potentials, and the $g_{\mu\nu}$.

---

† 'Ann. der Physik,' vol 79, p. 361 (1926).

‡ 'Roy. Soc. Proc.,' A, vol. 118, p. 654 (1928).

‖ *Ibid.*, vol. 117, p. 610 (1928); vol. 118, p. 351 (1928). See also Schlapp, *ibid.*, vol. 119, p. 313 (1928); Gordon, 'Z. f. Physik,' vol. 48, p. 11 (1928); Landé, *ibid.*, p. 601; Möglich, *ibid.*, p. 852; v. Neumann, *ibid.*, p. 868; and Breit, 'Proc. Nat. Acad. Science,' vol. 14, p. 553 (1928).

¶ 'Proc. Camb. Phil. Soc.,' vol. 24, p. 501 (1928).

** Darwin states (*loc cit.*, p. 657) that Dirac's equations can be thrown into tensor form if 16 wave functions are used. The 8 wave equations (4) of the second order are equivalent to the 16 equations (2)′, (3)′ of the first order and so may form a system such as Prof. Darwin has in mind.

§ 2. *The Wave Equations in a Null Gravitational Field.*

Let us start by supposing that if there is no external electromagnetic or gravitational field, each of the wave functions satisfies the equation

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \frac{4\pi^2 m^2 c^2}{h^2}\psi = 0. \tag{1}$$

The wave functions form two four-vectors (P_x, P_y, P_z, P_t) and (Q_x, Q_y, Q_z, Q_t), so that there will be eight of these equations. They can be written in the Maxwellian form

$$\left.\begin{aligned} -\frac{1}{c}\frac{\partial \mathbf{s}}{\partial t} - \operatorname{curl}\mathbf{b} &= \frac{2\pi i}{h}mc\mathbf{P} + \operatorname{grad} S \\ \operatorname{div}\mathbf{s} &= \frac{2\pi i}{h}mcP_t - \frac{1}{c}\frac{\partial S}{\partial t} \\ -\frac{1}{c}\frac{\partial \mathbf{b}}{\partial t} + \operatorname{curl}\mathbf{s} &= \frac{2\pi i}{h}mc\mathbf{Q} + \operatorname{grad} B \\ \operatorname{div}\mathbf{b} &= \frac{2\pi i}{h}mcQ_t - \frac{1}{c}\frac{\partial B}{\partial t} \end{aligned}\right\} \tag{2}$$

where $\mathbf{P}$, $\mathbf{Q}$ denote the space vectors (P_x, P_y, P_z), (Q_x, Q_y, Q_z), m is the rest mass of the electron, and the space vectors† $\mathbf{b}\,(b_x, b_y, b_z)$, $\mathbf{s}\,(s_x, s_y, s_z)$, and the scalars S, B are defined by

$$\left.\begin{aligned} \frac{2\pi i}{h}mc\mathbf{b} &= -\operatorname{curl}\mathbf{P} - \frac{1}{c}\frac{\partial \mathbf{Q}}{\partial t} - \operatorname{grad} Q_t \\ \frac{2\pi i}{h}mc\mathbf{s} &= \operatorname{curl}\mathbf{Q} - \frac{1}{c}\frac{\partial \mathbf{P}}{\partial t} - \operatorname{grad} P_t \\ \frac{2\pi i}{h}mcS &= \operatorname{div}\mathbf{P} + \frac{1}{c}\frac{\partial P_t}{\partial t} \\ \frac{2\pi i}{h}mcB &= \operatorname{div}\mathbf{Q} + \frac{1}{c}\frac{\partial Q_t}{\partial t} \end{aligned}\right\}. \tag{3}$$

If there is an external electromagnetic field of scalar potential V and vector potential $\mathbf{A}$ (A_1, A_2, A_3) the equations (2), (3) are modified in the familiar way

---

† In the former paper the eight wave functions were taken to be $\mathbf{b}$, $\mathbf{s}$, B, S, and $P_x \ldots Q_t$ were defined in terms of them by (2). But it is more convenient to take the four-vectors as fundamental.

$$\frac{1}{c}\frac{\partial}{\partial t} \text{ being replaced by } \frac{1}{c}\frac{\partial}{\partial t} - \frac{2\pi ie}{hc}V, \quad \frac{\partial}{\partial x} \text{ by } \frac{\partial}{\partial x} + \frac{2\pi ie}{hc}A_1,$$

etc. Thus, the first equation of (2) becomes

$$-\frac{1}{c}\frac{\partial \mathbf{s}}{\partial t} - \text{curl } \mathbf{b} - \text{grad } S + \frac{2\pi ie}{hc}\{V \,.\, \mathbf{s} - [\mathbf{A} \,.\, \mathbf{b}] - \mathbf{A} \,.\, S\} = \frac{2\pi i}{h} mc\mathbf{P}$$

Let the equations so modified be denoted by (2)′, (3)′. Then (2)′ are taken as the wave equations, the functions $\mathbf{b}$, $\mathbf{s}$, B, S being now defined by (3)′. On substituting from (3)′ in (2)′ the wave equations take the form

$$\left.\begin{aligned} &D\mathbf{P} + \frac{2\pi ie}{hc}\{-[\mathbf{H} \,.\, \mathbf{P}] + \mathbf{E} \,.\, P_t - [\mathbf{E} \,.\, \mathbf{Q}] - \mathbf{H} \,.\, Q_t\} = 0 \\ &DP_t + \frac{2\pi ie}{hc}\{(\mathbf{E} \,.\, \mathbf{P}) - (\mathbf{H} \,.\, \mathbf{Q})\} \qquad\qquad\qquad\quad = 0 \\ &D\mathbf{Q} + \frac{2\pi ie}{hc}\{-[\mathbf{H} \,.\, \mathbf{Q}] + \mathbf{E} \,.\, Q_t + [\mathbf{E} \,.\, \mathbf{P}] + \mathbf{H} \,.\, P_t\} = 0 \\ &DQ_t + \frac{2\pi ie}{hc}\{(\mathbf{E} \,.\, \mathbf{Q}) + (\mathbf{H} \,.\, \mathbf{P})\} \qquad\qquad\qquad\quad = 0 \end{aligned}\right\}, \tag{4}$$

where D is the operator

$$D = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{4\pi ie}{hc}\left(V\frac{1}{c}\frac{\partial}{\partial t} + A_1\frac{\partial}{\partial x} + A_2\frac{\partial}{\partial y} + A_3\frac{\partial}{\partial z}\right) - \frac{4\pi^2m^2c^2}{h^2} + \frac{4\pi^2e^2}{h^2c^2}(V^2 - A_1^2 - A_2^2 - A_3^2).$$

and $\mathbf{E}$, $\mathbf{H}$ are the electric and magnetic forces.

If we write

$$\left.\begin{aligned} &\psi_1 = P_z + iQ_t, \quad \psi_2 = P_x + iP_y, \quad \psi_3 = -P_t - iQ_z, \quad \psi_4 = Q_y - iQ_x \\ &\omega_1 = -P_x + iP_y, \quad \omega_2 = P_z - iQ_t, \quad \omega_3 = Q_y + iQ_x, \quad \omega_4 = P_t - iQ_z \end{aligned}\right\}, \tag{5}$$

the equations (4) are equivalent to

$$\left.\begin{aligned} &D\psi_1 + \frac{2\pi ie}{hc}(iH_x\psi_2 + H_y\psi_2 + iH_z\psi_1 - E_x\psi_4 + iE_y\psi_4 - E_z\psi_3) = 0 \\ &D\psi_2 + \frac{2\pi ie}{hc}(iH_x\psi_1 - H_y\psi_1 - iH_z\psi_2 - E_x\psi_3 - iE_y\psi_3 + E_z\psi_4) = 0 \\ &D\psi_3 + \frac{2\pi ie}{hc}(iH_x\psi_4 + H_y\psi_4 + iH_z\psi_3 - E_x\psi_2 + iE_y\psi_2 - E_z\psi_1) = 0 \\ &D\psi_4 + \frac{2\pi ie}{hc}(iH_x\psi_3 - H_y\psi_3 - iH_z\psi_4 - E_x\psi_1 - iE_y\psi_1 + E_z\psi_2) = 0 \end{aligned}\right\} \tag{6}$$

together with the same equations with the ω's in place of the ψ's. (6) are the second order equations satisfied by Dirac's functions $\psi_1, \psi_2, \psi_3, \psi_4$. They are given at the top of page 619 of the first of his memoirs cited above. If

$$\left.\begin{array}{l} \alpha_1 = -s_z + i\mathrm{B},\ \alpha_2 = -s_x - is_y,\ \alpha_3 = -ib_z + \mathrm{S},\ \alpha_4 = b_y - ib_x \\ \beta_1 = s_x - is_y,\ \beta_2 = -s_z - i\mathrm{B},\ \beta_3 = ib_x + b_y,\ \beta_4 = -ib_z - \mathrm{S} \end{array}\right\}, \tag{7}$$

the same equations hold in the α's and in the β's. In this notation (2)′, (3)′ are equivalent to four similar groups of four equations, namely, the equations

$$\begin{aligned} p_0\psi_1 + (p_1 - ip_2)\,\psi_4 + p_3\psi_3 &= -mc\alpha_1 \\ p_0\psi_2 + (p_1 + ip_2)\,\psi_3 - p_3\psi_4 &= -mc\alpha_2 \\ p_0\psi_3 + (p_1 - ip_2)\,\psi_2 + p_3\psi_1 &= mc\alpha_3 \\ p_0\psi_4 + (p_1 + ip_2)\,\psi_1 - p_3\psi_2 &= mc\alpha_4, \end{aligned}$$

where

$$p_0 = -\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} + \frac{e}{c}\mathrm{V},\quad p_1 = \frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{e}{c}\mathrm{A}_1, \text{ etc.},$$

together with three groups of similar equations. In the first group the α's are the ψ's change places Thus

$$p_0\alpha_1 + (p_1 - ip_2)\,\alpha_4 + p_3\alpha_3 = -mc\psi_1, \text{ etc.}$$

In the second group the ψ's are replaced by ω's and the α's by β's, and in the third group the ψ's are replaced by β's and the α's by ω's. Now Dirac's equations are the above with ψ's in place of α's, *i.e.*, the equations

$$p_0\psi_1 + (p_1 - ip_2)\,\psi_4 + p_3\psi_3 = -mc\psi_1, \text{ etc.}$$

If therefore $(\psi_1, \psi_2, \psi_3, \psi_4)$ is a solution of Dirac's equations, then a solution of (2)′, (3)′ is

$$\alpha_n = \beta_n = \omega_n = \psi_n,\quad (n = 1, 2, 3, 4). \tag{8}$$

Again, it will appear that the energy density and current are (save for a normalising constant)

$$\rho = (\mathbf{s}\,.\,\mathbf{P}^*) - (\mathbf{b}\,.\,\mathbf{Q}^*) + \mathrm{SP}_t^* - \mathrm{BQ}_t^* + \text{conjugate}$$

$$\mathbf{j} = c\,\{\mathbf{s}\mathrm{P}_t^* + [\mathbf{b}\,.\,\mathbf{P}^*] - \mathbf{b}\,.\,\mathrm{Q}_t^* + [\mathbf{s}\,.\,\mathbf{Q}^*] + \mathrm{S}\mathbf{P}^* - \mathrm{B}\mathbf{Q}^* + \text{conjugate}\}.$$

These can be written in the form

$$\rho = -\tfrac{1}{2}\,(\psi_1\alpha_1^* + \psi_2\alpha_2^* + \psi_3\alpha_3^* + \psi_4\alpha_4^* + \omega_1\beta_1^* + \omega_2\beta_2^* + \omega_3\beta_3^* + \omega_4\beta_4^* + \text{conjugate})$$

$$j_x = \frac{c}{2}\,(\psi_1\alpha_4^* + \psi_2\alpha_3^* + \psi_3\alpha_2^* + \psi_4\alpha_1^* + \omega_1\beta_4^* + \omega_2\beta_3^* + \omega_3\beta_2^* + \omega_4\beta_1^* + \text{conjugate}), \text{ etc.}$$

so that for a solution of the form (8),

$$\begin{aligned} \rho &= -2\,(\psi_1\psi_1^* + \psi_2\psi_2^* + \psi_3\psi_3^* + \psi_4\psi_4^*) \\ j_x &= 2c\,(\psi_1\psi_4^* + \psi_2\psi_3^* + \psi_3\psi_2^* + \psi_4\psi_1^*), \text{ etc.,} \end{aligned}$$

which are precisely the expressions found by Darwin[†] in connection with Dirac's theory. Thus the whole theory of the hydrogen atom worked out by Darwin and Schlapp is valid on the present theory.

§ 3. *The Principle of Least Action.*

The wave equations enable us to calculate the energy levels of a system, *e.g.*, the hydrogen atom in a magnetic field. The next step is to find the rules of combination which give the intensity and polarisation of the various lines in the spectrum. The wave equations do not suffice to determine these. They must be supplemented by equations expressing the energy density and current in terms of the wave functions.

These additional equations are derived by means of a principle of least action. The first step is to find an integral from which the wave equations and the complex conjugate equations can be deduced by the calculus of variations. The functions to be varied are the wave functions $P_x \ldots Q_t$ and their complex conjugates $P_x^* \ldots Q_t^*$. This process is similar to that by which Maxwell's equations

$$\left.\begin{aligned} \nabla^2 V - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} = 0 \\ \nabla^2 A_1 - \frac{1}{c^2}\frac{\partial^2 A_1}{\partial t^2} = 0, \text{ etc.} \end{aligned}\right\} \qquad (9)$$

are deduced from the integral

$$\tfrac{1}{2}\iiiint (H^2 - E^2)\, dx\, dy\, dz\, dt \qquad (10)$$

by varying the functions V, A_1, A_2, A_3. The atom and the radiation field are now combined into a single system by adding (10) to the integral found for the wave functions and then varying both sets of functions, $P_x \ldots Q_t^*$ and V, A_1, A_2, A_3.[‡]

It is possible to go a step further. An external gravitational field has not yet been contemplated, but it can be fitted into the scheme of equations in just the same way as the electromagnetic field. The integral corresponding to (10) is known to be

$$-\gamma \int (G - 2\lambda)\sqrt{-g}\, d\omega, \qquad (11)$$

[†] 'Roy. Soc. Proc.,' A, vol. 118, p. 660 (1928). Equations (3.3).

[‡] This process for combining two systems into one is, of course, classical. It has been applied to Dirac's theory by Darwin, *loc. cit.*

where G is the scalar curvature, $d\omega$ the element of co-ordinate hypervolume and γ, λ constants.† Einstein's extended field equations

$$G_{\mu\nu} = \lambda g_{\mu\nu}$$

are deduced from (11) by varying the gravitational potentials $g_{\mu\nu}$. If the wave integral and (10) are written in tensor form and added to (11) the three sets of equations, wave, electromagnetic, and gravitational, can be deduced by varying in turn the corresponding sets of functions. The sum of the three integrals is accordingly defined to be the action.

The integral from which the wave equations can be deduced is found as follows. Suppose first that there is no electromagnetic or gravitational field. Then the wave equations are very similar to (9) and the relation between the six-vector $(b_x, b_y, b_z, s_x, s_y, s_z)$ and the four-vectors **P**, **Q** resembles the relation between the six-vector $(E_x, E_y, E_z, H_x, H_y, H_z)$ and the four-vector (V, A_1, A_2, A_3). We infer that the integral required is similar to (10) and after a few trials the precise form is found to be

$$\iiiint\{(\mathbf{b} \,.\, \mathbf{b}^*) - (\mathbf{s} \,.\, \mathbf{s}^*) + SS^* - BB^* + (\mathbf{P} \,.\, \mathbf{P}^*) - (\mathbf{Q} \,.\, \mathbf{Q}^*) \\ - P_t \,.\, P_t^* + Q_t \,.\, Q_t^*\}\, dx\, dy\, dz\, dt, \quad (12)$$

f^* denoting, as usual, the complex conjugate of a function f. $b_x \ldots B^*$ are to be expressed in terms of $P_x \ldots Q_t^*$ by (3) and the complex conjugate equations. If there is an electromagnetic field the integral (12) is formally the same, but $b_x \ldots B^*$ are now to be expressed in terms of the wave functions by (3)′ and the complex conjugate equations.

The next step is to write the integral in tensor form. The wave functions are two four-vectors P_μ, Q_μ, and a six-vector $U_{\mu\nu}$ and two scalars S, B are defined in terms of them by the equations

$$\frac{2\pi i}{h} mc^2 U_{\mu\nu} = \frac{\partial P_\nu}{\partial x^\mu} - \frac{\partial P_\mu}{\partial x^\nu} + \sum_{\alpha,\beta} \frac{g_{\mu\alpha} g_{\nu\beta}}{\sqrt{-g}} \left(\frac{\partial Q_\delta}{\partial x^\gamma} - \frac{\partial Q_\gamma}{\partial x^\delta}\right) - \zeta_\mu P_\nu \\ + \zeta_\nu P_\mu - \sum_{\alpha,\beta} \frac{g_{\mu\alpha} g_{\nu\beta}}{\sqrt{-g}} (\zeta_\gamma Q_\delta - \zeta_\delta Q_\gamma) \quad (13)$$

$$\frac{2\pi i}{h} mc^2 S = (P^\alpha)_\alpha - \zeta_\alpha P^\alpha = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\alpha} (\sqrt{-g}\, P^\alpha) - \zeta_\alpha P^\alpha. \quad (14)$$

$$\frac{2\pi i}{h} mc^2 B = (Q^\alpha)_\alpha - \zeta_\alpha Q^\alpha, \quad (15)$$

---

† See Eddington, 'The Mathematical Theory of Relativity,' p. 137; E. T. Whittaker, 'Roy. Soc. Proc.,' A, vol. 113, p. 496 (1927) and references there given. The result is substantially due to Hilbert, 'Gött. Nach.,' p. 395 (1915).

where ϕ_μ is the electromagnetic potential, $\zeta_\mu = \frac{2\pi iec}{h}\phi_\mu$, and $\alpha\beta\gamma\delta$ is an even permutation of 0123. In regard to the first equation it should be remarked that if $A_{\alpha\beta}$ is a covariant skew tensor of the second rank, then $A_{\gamma\delta}$ is a contravariant tensor density. Thus

$$\frac{\partial Q_\delta}{\partial x^\gamma} - \frac{\partial Q_\gamma}{\partial x^\delta}$$

is a contravariant tensor density and

$$\frac{g_{\mu\alpha}\, g_{\nu\beta}}{\sqrt{-g}}\left(\frac{\partial Q_\delta}{\partial x^\gamma} - \frac{\partial Q_\gamma}{\partial x^\delta}\right)$$

is a covariant tensor.

If the metric is given by

$$ds^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2), \tag{16}$$

the components of the tensors are the functions previously introduced. Thus

$$\left.\begin{aligned} P_\mu &= \left(P_t, -\frac{1}{c}P_x, -\frac{1}{c}P_y, -\frac{1}{c}P_z\right) \\ Q_\mu &= \left(Q_t, -\frac{1}{c}Q_x, -\frac{1}{c}Q_y, -\frac{1}{c}Q_z\right) \\ \phi_\mu &= \left(\frac{1}{c}V, -\frac{1}{c^2}A_1, -\frac{1}{c^2}A_2, -\frac{1}{c^2}A_3\right) \\ U_{01} &= \frac{1}{c}s_x,\ U_{23} = \frac{1}{c^2}b_x, \quad \text{etc.} \end{aligned}\right\}. \tag{17}$$

and (13), (14), (15) become equations (3)′.

$$X_{\mu\nu} = \frac{\partial \phi_\nu}{\partial x^\mu} - \frac{\partial \phi_\mu}{\partial x^\nu} \tag{18}$$

is the electromagnetic force, and with the Galilean metric (16)

$$X_{01} = \frac{1}{c}E_x, \quad X_{23} = -\frac{1}{c^2}H_x, \quad \text{etc.}$$

The integrand of (12) is

$$D = \tfrac{1}{2}U^{\mu\nu}U_{\mu\nu}{}^* + SS^* - BB^* - P^\mu P_\mu{}^* + Q^\mu Q_\mu{}^*. \tag{19}$$

To this must be added the electromagnetic and gravitational terms, so that the complete "world-function" is

$$H = -\gamma(G - 2\lambda) + \tfrac{1}{4}X_{\mu\nu}X^{\mu\nu} + D. \tag{20}$$

The "principle of least action" states that

$$\delta \int \mathrm{H} \sqrt{-g}\, d\omega = 0, \tag{21}$$

wherein the integrand is regarded as a function of P_μ, $P_\mu{}^*$, Q_μ, $Q_\mu{}^*$, ϕ_μ, $g_{\mu\nu}$.

The variation of the first two terms has been studied by many previous writers, and discussions will be found in the works just referred to. It is found that

$$\begin{aligned}\delta \int \{-\gamma(G - 2\lambda) + \tfrac{1}{4} X_{\mu\nu} X^{\mu\nu}\} \sqrt{-g}\, d\omega \\ = \int \{\gamma(G^{\mu\nu} - \tfrac{1}{2} g^{\mu\nu} G + \lambda g^{\mu\nu}) + \tfrac{1}{2} E^{\mu\nu}\} \delta g_{\mu\nu} \sqrt{-g}\, d\omega \\ + \int \frac{\partial}{\partial x^\nu} (X^{\mu\nu} \sqrt{-g})\, \delta \phi_\mu\, d\omega,\end{aligned} \tag{22}$$

where

$$E^{\mu\nu} = \tfrac{1}{4} g^{\mu\nu} X_{\alpha\beta} X^{\alpha\beta} - X^{\mu\alpha} X^{\nu}{}_{\alpha}. \tag{23}$$

§ 4. *The Wave Equations in Tensor Form.*

There remains the term

$$\int D \sqrt{-g}\, d\omega,$$

which contains all the wave functions. Let us consider first the variation due to small changes in $P_\mu{}^*$, $Q_\mu{}^*$ while P_μ, Q_μ, ϕ_μ, $g_{\mu\nu}$ remain constant. For such variations

$$\begin{aligned}&\delta \int D \sqrt{-g}\, d\omega \\ &= \int \{\tfrac{1}{2} U^{\mu\nu} \delta U_{\mu\nu}{}^* + S\, \delta S^* - B\, \delta B^* - P^\mu\, \delta P_\mu{}^* + Q^\mu\, \delta Q_\mu{}^*\} \sqrt{-g}\, d\omega \\ &= \int \Big\{ -\tfrac{1}{2} \Big(\frac{h}{2\pi i m c^2}\Big) U^{\mu\nu}\, \delta \Big[\frac{\partial P_\nu{}^*}{\partial x^\mu} - \frac{\partial P_\mu{}^*}{\partial x^\nu} + \sum_{\alpha, \beta} \frac{g_{\mu\alpha} g_{\nu\beta}}{\sqrt{-g}} \Big(\frac{\partial Q_\delta{}^*}{\partial x^\gamma} - \frac{\partial Q_\gamma{}^*}{\partial x^\delta} \Big) \\ &\qquad + \zeta_\mu P_\nu{}^* - \zeta_\nu P_\mu{}^* + \sum_{\alpha, \beta} \frac{g_{\mu\alpha} g_{\nu\beta}}{\sqrt{-g}} (\zeta_\gamma Q_\delta{}^* - \zeta_\delta Q_\gamma{}^*) \Big] \\ &\qquad - \frac{h}{2\pi i m c^2} S \delta \Big[\frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\alpha} (\sqrt{-g}\, P^{*\alpha}) + \zeta_\alpha P^{*\alpha} \Big] \\ &\qquad + \frac{h}{2\pi i m c^2} B \delta \Big[\frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\alpha} (\sqrt{-g}\, Q^{*\alpha}) + \zeta_\alpha Q^{*\alpha} \Big] \\ &\qquad\qquad - P^\mu \delta P_\mu{}^* + Q^\mu\, \delta Q_\mu{}^* \Big\} \sqrt{-g}\, d\omega,\end{aligned}$$

on substituting from the equations conjugate to (13), (14), (15)

$$
\begin{aligned}
= \int\Big\{ & -\tfrac{1}{2}\Big(\frac{h}{2\pi imc^2}\Big)\mathrm{U}^{\mu\nu}\Big[\frac{\partial\delta\mathrm{P}_\nu{}^*}{\partial x^\mu} - \frac{\partial\delta\mathrm{P}_\mu{}^*}{\partial x^\nu} + \sum_{\alpha,\beta}\frac{g_{\mu\alpha}g_{\nu\beta}}{\sqrt{-g}}\Big(\frac{\partial\delta\mathrm{Q}_\delta{}^*}{\partial x^\gamma} - \frac{\partial\delta\mathrm{Q}_\gamma{}^*}{\partial x^\delta}\Big) \\
& + \zeta_\mu\,\delta\mathrm{P}_\nu{}^* - \zeta_\nu\,\delta\mathrm{P}_\mu{}^* + \sum_{\alpha,\beta}\frac{g_{\mu\alpha}g_{\nu\beta}}{\sqrt{-g}}(\zeta_\gamma\,\delta\mathrm{Q}_\delta{}^* - \zeta_\delta\,\delta\mathrm{Q}_\gamma{}^*)\Big] \\
& - \frac{h}{2\pi imc^2}\mathrm{S}\Big[\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\alpha}(\sqrt{-g}\,\delta\mathrm{P}^{*\alpha}) + \zeta_\alpha\,\delta\mathrm{P}^{*\alpha}\Big] \\
& + \frac{h}{2\pi imc^2}\mathrm{B}\Big[\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\alpha}(\sqrt{-g}\,\delta\mathrm{Q}^{*\alpha}) + \zeta_\alpha\,\delta\mathrm{Q}^{*\alpha}\Big] \\
& - \mathrm{P}^\mu\,\delta\mathrm{P}_\mu{}^* + \mathrm{Q}^\mu\,\delta\mathrm{Q}_\mu{}^*\Big\}\sqrt{-g}\,d\omega
\end{aligned}
$$

$$
\begin{aligned}
= \Big(-\frac{h}{2\pi imc^2}\Big)\int\Big\{ & \delta\mathrm{P}_\mu{}^*\Big[\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\nu}(\mathrm{U}^{\mu\nu}\sqrt{-g}) - \zeta_\nu\mathrm{U}^{\mu\nu} \\
& - g^{\mu\alpha}\frac{\partial\mathrm{S}}{\partial x^\alpha} + \zeta^\mu\mathrm{S} + \frac{2\pi imc^2}{h}\mathrm{P}^\mu\Big] + \delta\mathrm{Q}_\mu{}^*\Big[\frac{1}{\sqrt{-g}}\Big(\frac{\partial\mathrm{U}_{\kappa\lambda}}{\partial x^\nu} + \frac{\partial\mathrm{U}_{\lambda\nu}}{\partial x^\kappa} + \frac{\partial\mathrm{U}_{\nu\kappa}}{\partial x^\lambda}\Big) \\
& + g^{\mu\alpha}\frac{\partial\mathrm{B}}{\partial x^\alpha} - \frac{1}{\sqrt{-g}}(\zeta_\nu\mathrm{U}_{\kappa\lambda} + \zeta_\kappa\mathrm{U}_{\lambda\nu} + \zeta_\lambda\mathrm{U}_{\nu\kappa}) \\
& - \zeta^\mu\mathrm{B} - \frac{2\pi imc^2}{h}\mathrm{Q}^\mu\Big\}\sqrt{-g}\,d\omega,
\end{aligned}
$$

where $\kappa\lambda\mu\nu$ is an even permutation of 0123. The familiar process of integrating by parts and rejecting the integrated terms has been carried out. The coefficients of $\delta\mathrm{P}_\mu{}^*$, $\delta\mathrm{Q}_\mu{}^*$ can now be equated to zero. Thus

$$
\begin{aligned}
\frac{2\pi i}{h}mc^2\mathrm{P}^\mu &= -\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\nu}(\sqrt{-g}\,\mathrm{U}^{\mu\nu}) + g^{\mu\nu}\frac{\partial\mathrm{S}}{\partial x^\nu} + \zeta_\nu\mathrm{U}^{\mu\nu} - \zeta^\mu\mathrm{S} \\
&= -(\mathrm{U}^{\mu\nu})_\nu + g^{\mu\nu}\frac{\partial\mathrm{S}}{\partial x^\nu} + \zeta_\nu\mathrm{U}^{\mu\nu} - \zeta^\mu\mathrm{S}, \qquad (24)
\end{aligned}
$$

using the notation of covariant differentiation.

$$
\begin{aligned}
\frac{2\pi i}{h}mc^2\mathrm{Q}^\mu = \frac{1}{\sqrt{-g}}\Big(\frac{\partial\mathrm{U}_{\kappa\lambda}}{\partial x^\nu} + \frac{\partial\mathrm{U}_{\lambda\nu}}{\partial x^\kappa} + \frac{\partial\mathrm{U}_{\nu\kappa}}{\partial x^\lambda}\Big) + g^{\mu\nu}\frac{\partial\mathrm{B}}{\partial x^\nu} \\
- \frac{1}{\sqrt{-g}}(\zeta_\nu\mathrm{U}_{\kappa\lambda} + \zeta_\kappa\mathrm{U}_{\lambda\nu} + \zeta_\lambda\mathrm{U}_{\nu\kappa}) - \zeta^\mu\mathrm{B}. \qquad (25)
\end{aligned}
$$

In the Galilean case these reduce to the equations (2)′ If P_μ, Q_μ are varied instead of $\mathrm{P}_\mu{}^*$, $\mathrm{Q}_\mu{}^*$ (24), (25) are replaced by the complex conjugnte equations, *e.g.*,

$$
-\frac{2\pi imc^2}{h}\mathrm{P}^{*\mu} = -(\mathrm{U}^{*\mu\nu})_\nu + g^{\mu\nu}\frac{\partial\mathrm{S}^*}{\partial x^\nu} - \zeta_\nu\mathrm{U}^{*\mu\nu} + \zeta^\mu\mathrm{S}^*.
$$

The next step is the elimination of the expressions $U_{\mu\nu}$, S, B. Substitute in (24) from (13), (14), (15),

$$\begin{aligned}
&-\frac{4\pi^2m^2c^4}{h^2}\mathrm{P}^\mu \\
&\quad= -g^{\mu\sigma}g^{\nu\tau}\left[\frac{\partial \mathrm{P}_\tau}{\partial x^\sigma}-\frac{\partial \mathrm{P}_\sigma}{\partial x^\tau}+\sum_{\alpha,\beta}\frac{g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}\left(\frac{\partial \mathrm{Q}_\delta}{\partial x^\gamma}-\frac{\partial \mathrm{Q}_\gamma}{\partial x^\delta}\right)-\zeta_\sigma \mathrm{P}_\tau+\zeta_\tau \mathrm{P}_\sigma\right. \\
&\qquad\left.-\sum_{\alpha,\beta}\frac{g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}(\zeta_\gamma \mathrm{Q}_\delta-\zeta_\delta \mathrm{Q}_\gamma)\right]_\nu+g^{\mu\nu}[(\mathrm{P}^\alpha)_\alpha-\zeta_\alpha \mathrm{P}^\alpha]_\nu \\
&\qquad+g^{\mu\sigma}g^{\nu\tau}\zeta_\nu\left[\frac{\partial \mathrm{P}_\tau}{\partial x^\sigma}-\frac{\partial \mathrm{P}_\sigma}{\partial x^\tau}+\sum_{\alpha,\beta}\frac{g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}\left(\frac{\partial \mathrm{Q}_\delta}{\partial x^\gamma}-\frac{\partial \mathrm{Q}_\gamma}{\partial x^\delta}\right)-\zeta_\sigma \mathrm{P}_\tau+\zeta_\tau \mathrm{P}_\sigma\right. \\
&\qquad\left.-\sum_{\alpha,\beta}\frac{g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}(\zeta_\gamma \mathrm{Q}_\delta-\zeta_\delta \mathrm{Q}_\gamma)\right]-\zeta^\mu[(\mathrm{P}^\alpha)_\alpha-\zeta_\alpha \mathrm{P}^\alpha]. \qquad (26)
\end{aligned}$$

Now

$$\begin{aligned}
&-g^{\mu\sigma}g^{\nu\tau}\left[\sum_{\alpha,\beta}\frac{g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}\left(\frac{\partial \mathrm{Q}_\delta}{\partial x^\gamma}-\frac{\partial \mathrm{Q}_\gamma}{\partial x^\delta}\right)\right]_\nu \\
&\quad= -\left[\sum_{\alpha,\beta}\frac{g^{\mu\sigma}g^{\nu\tau}g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}\left(\frac{\partial \mathrm{Q}_\delta}{\partial x^\gamma}-\frac{\partial \mathrm{Q}_\gamma}{\partial x^\delta}\right)\right]_\nu \\
&\quad= -\sum_\nu\left[\frac{1}{\sqrt{-g}}\left(\frac{\partial \mathrm{Q}_\lambda}{\partial x^\kappa}-\frac{\partial \mathrm{Q}_\kappa}{\partial x^\lambda}\right)\right]_\nu.
\end{aligned}$$

and it has been remarked above that the expression in square brackets is a contravariant skew tensor of rank two, say $V^{\mu\nu}$. Thus

$$\begin{aligned}
-\sum_\nu\left[\frac{1}{\sqrt{-g}}\left(\frac{\partial \mathrm{Q}_\lambda}{\partial x^\kappa}-\frac{\partial \mathrm{Q}_\kappa}{\partial x^\lambda}\right)\right]_\nu &= -(\mathrm{V}^{\mu\nu})_\nu = -\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\nu}(\mathrm{V}^{\mu\nu}\sqrt{-g}) \\
&= -\frac{1}{\sqrt{-g}}\sum_\nu\frac{\partial}{\partial x^\nu}\left(\frac{\partial \mathrm{Q}_\lambda}{\partial x^\kappa}-\frac{\partial \mathrm{Q}_\kappa}{\partial x^\lambda}\right)=0.
\end{aligned}$$

In the same way

$$g^{\mu\sigma}g^{\nu\tau}\left[\sum_{\alpha,\beta}\frac{g_{\sigma\alpha}g_{\tau\beta}}{\sqrt{-g}}(\zeta_\gamma \mathrm{Q}_\delta-\zeta_\delta \mathrm{Q}_\gamma)\right]_\nu=\frac{1}{\sqrt{-g}}\sum_\nu\frac{\partial}{\partial x^\nu}(\zeta_\kappa \mathrm{Q}_\lambda-\zeta_\lambda \mathrm{Q}_\kappa).$$

Thus in (26)

$$\begin{aligned}
&-\frac{4\pi^2m^2c^4}{h^2}\mathrm{P}^\mu \\
&\quad= -g^{\mu\sigma}g^{\nu\tau}(\mathrm{P}_{\tau\sigma\nu}-\mathrm{P}_{\sigma\tau\nu})+(\zeta^\mu \mathrm{P}^\nu)_\nu-(\zeta^\nu \mathrm{P}^\mu)_\nu \\
&\qquad\qquad+\frac{1}{\sqrt{-g}}\sum_\nu\frac{\partial}{\partial x^\nu}(\zeta_\kappa \mathrm{Q}_\lambda-\zeta_\lambda \mathrm{Q}_\kappa) \\
&\qquad+g^{\mu\nu}(\mathrm{P}^\alpha)_{\alpha\nu}-g^{\mu\nu}(\zeta_\alpha \mathrm{P}^\alpha)_\nu+\zeta^\tau g^{\mu\sigma}(\mathrm{P}_{\tau\sigma}-\mathrm{P}_{\sigma\tau}) \\
&\qquad+\frac{1}{\sqrt{-g}}\sum_\nu\zeta_\nu\left(\frac{\partial \mathrm{Q}_\lambda}{\partial x^\kappa}-\frac{\partial \mathrm{Q}_\kappa}{\partial x^\lambda}\right)-\zeta_\nu(\zeta^\mu \mathrm{P}^\nu-\zeta^\nu \mathrm{P}^\mu) \\
&\qquad\qquad-\frac{1}{\sqrt{-g}}\sum_\nu\zeta_\nu(\zeta_\kappa \mathrm{Q}_\lambda-\zeta_\lambda \mathrm{Q}_\kappa) \\
&\qquad-\zeta^\mu(\mathrm{P}^\alpha)_\alpha+\zeta^\mu\zeta_\alpha \mathrm{P}^\alpha
\end{aligned}$$

$$
\begin{aligned}
&= g^{\tau\nu}(P^\mu)_{\tau\nu} - g^{\mu\sigma}g^{\nu\tau}(P_{\tau\nu\sigma} - P_\epsilon B^\epsilon{}_{\tau\nu\sigma}) + (\zeta^\mu)_\nu P^\nu + \zeta^\mu (P^\nu)_\nu \\
&\quad - (\zeta^\nu)_\nu P^\mu - \zeta^\nu (P^\mu)_\nu + \frac{1}{\sqrt{-g}} \sum_\nu \left(\frac{\partial \zeta_\kappa}{\partial x^\nu} Q_\lambda - \frac{\partial \zeta_\lambda}{\partial x^\nu} Q_\kappa\right) \\
&\quad + \frac{1}{\sqrt{-g}} \sum_\nu \left(\zeta_\kappa \frac{\partial Q_\lambda}{\partial x^\nu} - \zeta_\lambda \frac{\partial Q_\kappa}{\partial x^\nu}\right) \\
&\qquad\qquad + g^{\mu\nu}(P^\alpha)_{\alpha\nu} - g^{\mu\nu}\zeta_\alpha (P^\alpha)_\nu - g^{\mu\nu}\zeta_{\alpha\nu} P^\alpha \\
&\quad + \zeta^\tau g^{\mu\sigma} P_{\tau\sigma} - \zeta^\tau (P^\mu)_\tau + \frac{1}{\sqrt{-g}} \sum_\nu \zeta_\nu \left(\frac{\partial Q_\lambda}{\partial x^\kappa} - \frac{\partial Q_\kappa}{\partial x^\lambda}\right) - \zeta_\nu \zeta^\mu P^\nu \\
&\quad + \zeta_\nu \zeta^\nu P^\mu - \frac{1}{\sqrt{-g}} \sum_\nu \zeta_\nu (\zeta_\kappa Q_\lambda - \zeta_\lambda Q_\kappa) - \zeta^\mu (P^\alpha)_\alpha + \zeta^\mu \zeta_\alpha P^\alpha .
\end{aligned}
$$

After some cancelling and rearrangement, using in particular the fact that

$$
(\zeta^\nu)_\nu = \frac{2\pi i e c}{h} (\phi^\nu)_\nu = 0, \tag{27}
$$

this leads to

$$
\begin{aligned}
&- \frac{4\pi^2 m^2 c^4}{h^2} P^\mu , \\
&\quad = g^{\tau\nu}(P^\mu)_{\tau\nu} - g^{\mu\sigma}g^{\nu\tau}P_{\tau\nu\sigma} + P_\epsilon g^{\mu\sigma} g^{\nu\tau} B^\epsilon{}_{\tau\nu\sigma} + (\zeta^\mu)_\nu P^\nu - \zeta^\nu (P^\mu)_\nu \\
&\qquad + \frac{1}{\sqrt{-g}} \sum_\nu Q_\nu \left(\frac{\partial \zeta_\lambda}{\partial x^\kappa} - \frac{\partial \zeta_\kappa}{\partial x^\lambda}\right) + g^{\mu\nu}(P^\alpha)_{\alpha\nu} - g^{\mu\nu}\zeta_\alpha (P^\alpha)_\nu - g^{\mu\nu}\zeta_{\alpha\nu} P^\alpha \\
&\qquad + \zeta^\tau g^{\mu\sigma} P_{\tau\sigma} - \zeta^\tau (P^\mu)_\tau + \zeta_\nu \zeta^\nu P^\mu \\
&\quad = g^{\tau\nu}(P^\mu)_{\tau\nu} + G^{\mu\epsilon} P_\epsilon - 2\zeta^\nu (P^\mu)_\nu + \frac{1}{\sqrt{-g}} \sum_\nu Q_\nu \left(\frac{\partial \zeta_\lambda}{\partial x^\kappa} - \frac{\partial \zeta_\kappa}{\partial x^\lambda}\right) \\
&\qquad + P^\nu \{(\zeta^\mu)_\nu - g^{\mu\alpha}\zeta_{\nu\alpha}\} + \zeta_\nu \zeta^\nu P^\mu ,
\end{aligned}
$$

or

$$
\begin{aligned}
g^{\sigma\tau}(P^\mu)_{\sigma\tau} &+ G^{\mu\nu}P_\nu - \frac{4\pi i e c}{h} \phi^\nu (P^\mu)_\nu + \frac{4\pi^2 m^2 c^4}{h^2}\left(1 - \frac{e^2}{m^2 c^2}\phi_\nu \phi^\nu\right) P^\mu \\
&- \frac{2\pi i e c}{h}\left\{X^{\mu\nu}P_\nu - \frac{1}{\sqrt{-g}}(X_{\kappa\lambda}Q_\nu + X_{\lambda\nu}Q_\kappa + X_{\nu\kappa}Q_\lambda)\right\} = 0.
\end{aligned} \tag{28}
$$

This is the tensor form of the first four of the eight wave equations (4). The corresponding equation in Q^μ is found as follows. Let

$$
\sqrt{-g}\, U^{\mu\nu} = T_{\kappa\lambda}
$$

where, as always, $\kappa\lambda\mu\nu$ is an even permutation of 0123. Then by Jacobi's theorem on the minors of the adjugate determinant,

$$
g\,(g^{\kappa\gamma}g^{\lambda\delta} - g^{\kappa\delta}g^{\lambda\gamma}) = g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha}, \tag{29}
$$

so that

$$\begin{aligned}\frac{2\pi i}{h} mc^2 T_{\mu\nu} &= \frac{\partial Q_\nu}{\partial x^\mu} - \frac{\partial Q_\mu}{\partial x^\nu} + \sqrt{-g}\, g^{\kappa\gamma} g^{\lambda\delta} \left(\frac{\partial P_\delta}{\partial x^\gamma} - \frac{\partial P_\gamma}{\partial x^\delta}\right) + \text{terms in } \phi_\mu \\ &= \frac{\partial Q_\nu}{\partial x^\mu} - \frac{\partial Q_\mu}{\partial x^\nu} + \tfrac{1}{2}\sqrt{-g}\,(g^{\kappa\gamma} g^{\lambda\delta} - g^{\kappa\delta} g^{\lambda\gamma}) \left(\frac{\partial P_\delta}{\partial x^\gamma} - \frac{\partial P_\gamma}{\partial x^\delta}\right) + \text{terms in } \phi_\mu \\ &= \frac{\partial Q_\nu}{\partial x^\mu} - \frac{\partial Q_\mu}{\partial x^\nu} - \tfrac{1}{2} \sum_{\alpha,\beta} \frac{g_{\mu\alpha} g_{\nu\beta} - g_{\mu\beta} g_{\nu\alpha}}{\sqrt{-g}} \left(\frac{\partial P_\delta}{\partial x^\gamma} - \frac{\partial P_\gamma}{\partial x^\delta}\right) + \text{terms in } \phi_\mu \\ &= \frac{\partial Q_\nu}{\partial x^\mu} - \frac{\partial Q_\mu}{\partial x^\nu} - \sum_{\alpha,\beta} \frac{g_{\mu\alpha} g_{\nu\beta}}{\sqrt{-g}} \left(\frac{\partial P_\delta}{\partial x^\gamma} - \frac{\partial P_\gamma}{\partial x^\delta}\right) - \zeta_\mu Q_\nu \\ &\qquad + \zeta_\nu Q_\mu + \sum_{\alpha,\beta} \frac{g_{\mu\alpha} g_{\nu\beta}}{\sqrt{-g}} (\zeta_\gamma P_\delta - \zeta_\delta P_\gamma), \qquad (30)\end{aligned}$$

also

$$\begin{aligned}U_{\kappa\lambda} = g_{\kappa\alpha} g_{\lambda\beta} U^{\alpha\beta} &= g_{\kappa\alpha} g_{\lambda\beta} \frac{1}{\sqrt{-g}} T_{\gamma\delta} = \tfrac{1}{2} (g_{\kappa\alpha} g_{\lambda\beta} - g_{\kappa\beta} g_{\lambda\alpha}) \frac{1}{\sqrt{-g}} T_{\gamma\delta} \\ &= \tfrac{1}{2} g\, (g^{\mu\gamma} g^{\nu\delta} - g^{\mu\delta} g^{\nu\gamma}) \frac{1}{\sqrt{-g}} T_{\gamma\delta} = -\sqrt{-g}\, g^{\mu\gamma} g^{\nu\delta} T_{\gamma\delta} = -\sqrt{-g}\, T^{\mu\nu},\end{aligned}$$

so that (25) can be written

$$\frac{2\pi i}{h} mc^2 Q^\mu = -(T^{\mu\nu})_\nu + g^{\mu\nu} \frac{\partial B}{\partial x^\nu} + \zeta_\nu T^{\mu\nu} - \zeta^\mu B. \qquad (31)$$

If $T_{\mu\nu}$ is written for $U_{\mu\nu}$, B for S, Q for P, and $-P$ for Q, (13), (14), (24) become respectively (30), (15), (31). Thus the final equation for Q^μ is

$$\begin{aligned}g^{\sigma\tau} (Q^\mu)_{\sigma\tau} + G^{\mu\nu} Q_\nu - \frac{4\pi iec}{h} \phi^\nu (Q^\mu)_\nu + \frac{4\pi^2 m^2 c^4}{h^2} \left(1 - \frac{e^2}{m^2 c^2} \phi_\nu \phi^\nu\right) Q^\mu \\ - \frac{2\pi iec}{h} \left\{ X^{\mu\nu} Q_\nu + \frac{1}{\sqrt{-g}} (X_{\kappa\lambda} P_\nu + X_{\lambda\nu} P_\kappa + X_{\nu\kappa} P_\lambda) \right\} = 0, \quad (32)\end{aligned}$$

(28), (32) are the wave equations in a gravitational field. It is to be observed that owing to the presence of the terms $G^{\mu\nu} P_\nu$, $G^{\mu\nu} Q_\nu$ they cannot be obtained by simply rewriting (4) in tensor form.

§ 5. *The Electromagnetic Equations.*

The next step consists in varying ϕ_μ in the wave part of the action, keeping $g_{\mu\nu}$ and the wave functions fixed. For such variations

$$\delta \int D\sqrt{-g}\, d\omega$$

$$= -\frac{e}{mc}\int\Big\{\tfrac{1}{2}U^{\mu\nu}\Big[\delta\phi_\mu P_\nu^* - \delta\phi_\nu P_\mu^* + \sum_{\alpha\,\beta}\frac{g_{\mu\alpha}g_{\nu\beta}}{\sqrt{-g}}(\delta\phi_\gamma Q_\delta^* - \delta\phi_\delta Q_\gamma^*)\Big]$$

$$+ \tfrac{1}{2}U^{*\mu\nu}\Big[\delta\phi_\mu P_\nu - \delta\phi_\nu P_\mu + \sum_{\alpha\,\beta}\frac{g_{\mu\alpha}g_{\nu\beta}}{\sqrt{-g}}(\delta\phi_\gamma Q_\delta - \delta\phi_\delta Q_\gamma)\Big]$$

$$+ SP^{*\alpha}\delta\phi_\alpha - BQ^{*\alpha}\delta\phi_\alpha + S^*P^\alpha\delta\phi_\alpha - B^*Q^\alpha\delta\phi_\alpha\Big\}\sqrt{-g}\, d\omega$$

$$= \frac{e}{mc}\int\Big\{-U^{\mu\nu}P_\nu^* - \frac{1}{\sqrt{-g}}(U_{\kappa\lambda}Q_\nu^* + U_{\lambda\nu}Q_\kappa^* + U_{\nu\kappa}Q_\lambda^*)$$

$$- SP^{*\mu} + BQ^{*\mu} + \text{conjugate}\Big\}\,\delta\phi_\mu\sqrt{-g}\, d\omega$$

The term in $\delta\phi_\mu$ contributed by the remainder of the action has been found in § 3, equation (22). The condition that the complete coefficient is zero yields the equation

$$\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\nu}(X^{\mu\nu}\sqrt{-g}) = \frac{e}{mc}\Big\{U^{\mu\nu}P_\nu^* + \frac{1}{\sqrt{-g}}(U_{\kappa\lambda}Q_\nu^* + U_{\lambda\nu}Q_\kappa^* + U_{\nu\kappa}Q_\lambda^*)$$

$$+ SP^{*\mu} - BQ^{*\mu} + \text{conjugate}\Big\}. \qquad (33)$$

The expression on the right must therefore be interpreted as the electromagnetic current vector. Thus

$$J^\mu = \frac{e}{mc}(U^{\mu\nu}P_\nu^* - T^{\mu\nu}Q_\nu^* + SP^{*\mu} - BQ^{*\mu} + \text{conjugate}). \qquad (34)$$

With the Galilean metric (16) the components are

$$\left.\begin{aligned} \rho = J^0 &= \frac{e}{mc}\{(\mathbf{s}\,.\,\mathbf{P}^*) - (\mathbf{b}\,.\,\mathbf{Q}^*) + S\,.\,P_t^* - B\,.\,Q_t^* + \text{conjugate}\} \\ \mathbf{j} = (J^1, J^2, J^3) &= \frac{e}{m}\{\mathbf{s}P_t^* + [\mathbf{b}\,.\,\mathbf{P}^*] - \mathbf{b}\,.\,Q_t^* + [\mathbf{s}\,.\,\mathbf{Q}^*] \\ &\qquad + S\mathbf{P}^* - B\mathbf{Q}^* + \text{conjugate}\} \end{aligned}\right\} \qquad (35)$$

The conservation equation

$$(J^\mu)_\mu = 0$$

is satisfied in virtue of the wave equations, as may be verified by straightforward analysis.

§ 6. *The Gravitational Equations.*

Finally, the field equations of gravitation are found by equating to zero the coefficient of $\delta g_{\mu\nu}$. The latter, a rather complicated expression, is found as follows. We have

$$\sqrt{-g}\,\mathrm{U}^{\mu\nu}\mathrm{U}_{\mu\nu}{}^{*} = \sum_{\kappa,\lambda} \mathrm{T}_{\kappa\lambda}\mathrm{U}_{\mu\nu}{}^{*}$$

$$= \frac{1}{\sqrt{-g}} \sum_{\kappa,\lambda}\sum_{\alpha,\beta} g_{\kappa\alpha} g_{\lambda\beta}\,.\, g\mathrm{L}^{\kappa\alpha\lambda\beta} + \text{terms not involving } g_{\mu\nu},$$

where

$$g\mathrm{L}^{\kappa\alpha\lambda\beta} = \frac{h^2}{4\pi^2 m^2 c^4}\Big\{-\Big(\frac{\partial \mathrm{P}_\delta}{\partial x^\gamma} - \frac{\partial \mathrm{P}_\gamma}{\partial x^\delta} - \zeta_\gamma \mathrm{P}_\delta + \zeta_\delta \mathrm{P}_\gamma\Big)\Big(\frac{\partial \mathrm{P}_\nu{}^*}{\partial x^\mu} - \frac{\partial \mathrm{P}_\mu{}^*}{\partial x^\nu} + \zeta_\mu \mathrm{P}_\nu{}^* - \zeta_\nu \mathrm{P}_\mu{}^*\Big)$$
$$+ \Big(\frac{\partial \mathrm{Q}_\nu}{\partial x^\mu} - \frac{\partial \mathrm{Q}_\mu}{\partial x^\nu} - \zeta_\mu \mathrm{Q}_\nu + \zeta_\nu \mathrm{Q}_\mu\Big)\Big(\frac{\partial \mathrm{Q}_\delta{}^*}{\partial x^\gamma} - \frac{\partial \mathrm{Q}_\gamma{}^*}{\partial x^\delta} + \zeta_\gamma \mathrm{Q}_\delta{}^* - \zeta_\delta \mathrm{Q}_\gamma{}^*\Big)\Big\}. \tag{36}$$

Now alter the dummy suffixes, supposing μξπρ, νστυ to be even permutations of 0123. Thus

$$\sqrt{-g}\,\mathrm{U}^{\mu\nu}\mathrm{U}_{\mu\nu}{}^{*} = \frac{1}{\sqrt{-g}} \sum_{\mu,\xi}\sum_{\nu,\sigma} g_{\mu\nu} g_{\xi\sigma}\cdot g\mathrm{L}^{\mu\nu\xi\sigma} + \text{terms not involving } g_{\mu\nu},$$

and the variation in this expression due to changes in $g_{\mu\nu}$ is

$$\delta(\sqrt{-g}\,\mathrm{U}^{\mu\nu}\mathrm{U}_{\mu\nu}{}^{*}) = -2\sqrt{-g}\,\mathrm{L}^{\mu\nu}\,\delta g_{\mu\nu} + \tfrac{1}{2}\sqrt{-g}\,\mathrm{L}g^{\mu\nu}\,\delta g_{\mu\nu},$$

where

$$\mathrm{L}^{\mu\nu} = g_{\xi\sigma}\mathrm{L}^{\mu\nu\xi\sigma}, \qquad \mathrm{L} = g_{\mu\nu}\mathrm{L}^{\mu\nu}.$$

Again, it is readily shown that

$$\delta\int \mathrm{SS}^*\sqrt{-g}\,d\omega = \int \mathrm{N}^{\mu\nu}\sqrt{-g}\,\delta g_{\mu\nu}\,d\omega,$$

$$\delta\int \mathrm{BB}^*\sqrt{-g}\,d\omega = \int \mathrm{O}^{\mu\nu}\sqrt{-g}\,\delta g_{\mu\nu}\,d\omega,$$

where

$$\mathrm{N}^{\mu\nu} = (\tfrac{1}{2}\mathrm{P}^{*\alpha}g^{\mu\nu} - \mathrm{P}^{*\nu}g^{\mu\alpha})\Big(\frac{h}{2\pi i m c^2}\frac{\partial \mathrm{S}}{\partial x^\alpha} - \frac{e}{mc}\phi_\alpha \mathrm{S}\Big) + \text{conjugate} - \tfrac{1}{2}\mathrm{SS}^* g^{\mu\nu}, \tag{37}$$

$$\mathrm{O}^{\mu\nu} = (\tfrac{1}{2}\mathrm{Q}^{*\alpha}g^{\mu\nu} - \mathrm{Q}^{*\nu}g^{\mu\alpha})\Big(\frac{h}{2\pi i m c^2}\frac{\partial \mathrm{B}}{\partial x^\alpha} - \frac{e}{mc}\phi_\alpha \mathrm{B}\Big) + \text{conjugate} - \tfrac{1}{2}\mathrm{BB}^* g^{\mu\nu}, \tag{38}$$

and that

$$\delta\int \mathrm{P}^\mu \mathrm{P}_\mu{}^*\sqrt{-g}\,d\omega = \int(-\mathrm{P}^\mu \mathrm{P}^{*\nu} + \tfrac{1}{2}\mathrm{P}^\alpha \mathrm{P}_\alpha{}^* g^{\mu\nu})\sqrt{-g}\,\delta g_{\mu\nu}\,d\omega.$$

The coefficient of $\sqrt{-g}\,\delta g_{\mu\nu}$ in the integrand of

$$\delta\int \mathrm{H}\sqrt{-g}\,d\omega$$

can now be equated to zero, and using the above results and (22) it is found that

$$\gamma\,(\mathrm{G}^{\mu\nu} - \tfrac{1}{2}\mathrm{G}g^{\mu\nu} + \lambda g^{\mu\nu}) + \tfrac{1}{2}\mathrm{E}^{\mu\nu} + \tfrac{1}{2}\mathrm{M}^{\mu\nu} = 0, \tag{39}$$

where

$$\begin{aligned}\mathrm{M}^{\mu\nu} = {} & -2\mathrm{L}^{\mu\nu} + (\tfrac{1}{2}\mathrm{L} - \mathrm{P}^{\alpha}\mathrm{P}_{\alpha}{}^{*} + \mathrm{Q}^{\alpha}\mathrm{Q}_{\alpha}{}^{*})\,g^{\mu\nu} + 2\mathrm{N}^{\mu\nu} \\ & - 2\mathrm{O}^{\mu\nu} + 2\mathrm{P}^{\mu}\mathrm{P}^{*\nu} - 2\mathrm{Q}^{\mu}\mathrm{Q}^{*\nu}.\end{aligned} \tag{40}$$

$\mathrm{E}^{\mu\nu}$ is the electromagnetic energy tensor and $\mathrm{M}^{\mu\nu}$ is interpreted as the energy tensor due to matter. The complete scheme has now been derived. It consists of the wave equations (28), (32), the Maxwell equations (33), and the Einstein equations (39).

It is a consequence of a general theorem due to Emmy Noether[†] that four of these equations must be a consequence of the others: in fact that four identities similar to the identities

$$\operatorname{div}\,(\mathrm{G}_{\mu}{}^{\nu} - \tfrac{1}{2}\mathrm{G}g_{\mu}{}^{\nu}) = 0$$

must exist.

The process of finding these identities is straightforward and the details may be omitted. The result is as follows. Let

$$\begin{aligned}\tfrac{1}{2}\mathrm{A}^{\mu\nu} &= \gamma\,(\mathrm{G}^{\mu\nu} - \tfrac{1}{2}\mathrm{G}g^{\mu\nu} + \lambda g^{\mu\nu}) + \tfrac{1}{2}\mathrm{E}^{\mu\nu} + \tfrac{1}{2}\mathrm{M}^{\mu\nu} \\ \mathrm{B}^{\mu} &= \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^{\nu}}(\mathrm{X}^{\mu\nu}\sqrt{-g}) - \mathrm{J}^{\mu} = (\mathrm{X}^{\mu\nu})_{\nu} - \mathrm{J}^{\mu}, \\ \mathrm{C}^{\mu} &= \frac{h}{2\pi i mc^{2}}\left\{(\mathrm{U}^{\mu\nu})_{\nu} - g^{\mu\nu}\frac{\partial \mathrm{S}}{\partial x^{\nu}} - \zeta_{\nu}\mathrm{U}^{\mu\nu} + \zeta^{\mu}\mathrm{S}\right\} + \mathrm{P}^{\mu}, \\ \mathrm{D}^{\mu} &= \frac{h}{2\pi i mc^{2}}\left\{(\mathrm{T}^{\mu\nu})_{\nu} - g^{\mu\nu}\frac{\partial \mathrm{B}}{\partial x^{\nu}} - \zeta_{\nu}\mathrm{T}^{\mu\nu} + \zeta^{\mu}\mathrm{B}\right\} + \mathrm{Q}^{\mu}.\end{aligned}$$

Then identically

$$\begin{aligned}(\mathrm{A}_{\mu}{}^{\nu})_{\nu} - \mathrm{B}^{\nu}\mathrm{X}_{\mu\nu} + \mathrm{C}^{\nu}\left(\frac{\partial \mathrm{P}_{\nu}{}^{*}}{\partial x^{\mu}} - \frac{\partial \mathrm{P}_{\mu}{}^{*}}{\partial x^{\nu}}\right) - \mathrm{P}_{\mu}{}^{*}\,(\mathrm{C}^{\nu})_{\nu} - \mathrm{D}^{\nu}\left(\frac{\partial \mathrm{Q}_{\nu}{}^{*}}{\partial x^{\mu}} - \frac{\partial \mathrm{Q}_{\mu}{}^{*}}{\partial x^{\nu}}\right) \\ + \mathrm{Q}_{\mu}{}^{*}\,(\mathrm{D}^{\nu})_{\nu} + \mathrm{C}^{*\nu}\left(\frac{\partial \mathrm{P}_{\nu}}{\partial x^{\mu}} - \frac{\partial \mathrm{P}_{\mu}}{\partial x^{\nu}}\right) - \mathrm{P}_{\mu}\,(\mathrm{C}^{*\nu})_{\nu} - \mathrm{D}^{*\nu}\left(\frac{\partial \mathrm{Q}_{\nu}}{\partial x^{\mu}} - \frac{\partial \mathrm{Q}_{\mu}}{\partial x^{\nu}}\right) \\ + \mathrm{Q}_{\mu}\,(\mathrm{D}^{*\nu})_{\nu} = 0.\end{aligned}$$

Thus the electromagnetic equations

$$\mathrm{B}^{\mu} = 0$$

are a consequence of the remaining equations

$$\mathrm{A}^{\mu\nu} = 0,\quad \mathrm{C}^{\mu} = 0,\quad \mathrm{D}^{\mu} = 0,\quad \mathrm{C}^{*\mu} = 0,\quad \mathrm{D}^{*\mu} = 0.$$

[†] 'Gött. Nach.,' p. 236 (1918). See § 5 of Prof. Whittaker's memoir "On Hilbert's World-Function," 'Roy. Soc. Proc.,' A, vol. 113, p. 496 (1927).

The Strength of Tubular Struts.

By Andrew Robertson, D.Sc., University of Bristol.

(Communicated by R. V. Southwell, F.R.S.—Received May 30, 1928.)

[Plates 10–14.]

Introduction.

The problem of the strength of a circular tubular strut differs from that of a solid strut of the same material in that there may be a condition of instability in the wall of the tube apart altogether from that type of instability considered in the Eulerian analysis which involves only the bending of the centre line of the strut. It is found that when a thin tube is tested in direct compression up to collapse, it fails by the formation of characteristic folds in the wall, examples of which are given in the photographs, figs. 7, 10 and 11 (Plate 11). The folds are of two types depending upon the ratio of the thickness t to the radius R of the tube. Above a certain value of t/R the fold consists of a uniform bulge in which all cross sections remain circular, below this value of t/R the walls "cave in" in several places producing a series of lobes. The phenomenon has been called wrinkling, and also secondary flexure, to distinguish it from the flexure produced by the bending of the centre line of the strut.

The problem has been investigated mathematically and experimentally. On the mathematical side the best-known investigations are those of Lilly, Southwell and Dean; on the experimental side those of Lilly, Mason, Barling and Webb, and Popplewell and Carrington.[*] The present investigation is experimental and the greater part of it was carried out in 1915, at which date the only investigations with which the author was acquainted were those of Lilly and Mason. A paper describing the work was presented in 1915 to a Sub-Committee of the War Committee of the Royal Society and was also communicated to the Air Departments of the Army and Navy. General publication, however, was withheld until the end of the war. A short summary of the conclusions reached was given to the British Association Meeting in 1919. Since the major part of the work was completed and reported upon, experimental investigations have been published by Messrs. Barling and Webb and Messrs.

---

[*] Lilly, 'Inst. Civ. Eng. Irel.,' March, 1906; Southwell, 'Phil. Trans.,' A, vol. 213, p. 187 (1914); Dean, 'Roy. Soc. Proc.,' A, vol. 107, p. 734 (1925); Mason, 'Proc. Inst. Mech. Eng.,' parts 3, 4 (1909); Barling and Webb, 'J. R. Aeron. Soc.,' Oct. (1918); Popplewell and Carrington, 'Proc. Inst. Civ. Eng.,' vol. 203, p. 381 (1917).

Popplewell and Carrington. As these results differ notably from the author's, some additional experiments were carried out on large tubes made from high tensile steel strip and the whole of the work is now presented in this paper. It will be convenient to depart from the strict chronological order in which the various experiments were made and examine the whole problem in the light of all the existing work known to the author.

Other Investigations.

Theoretical.—On the theoretical side Lilly obtained formulæ for the collapsing load of a tube and for the length of the characteristic wave-like fold produced when failure occurred by the formation of a uniform bulge. The formulæ are

$$p = \frac{E}{\sqrt{3}}\frac{t}{R}, \qquad \lambda = \frac{\pi}{4\sqrt{12}}\sqrt{Rt},$$

where p = collapsing load per square inch, E = Young's modulus, t = thickness, R = radius, λ = wave-length.

Southwell has investigated the same problem in his paper on the general theory of elastic stability and arrived at a formula identical with that of Lilly if the term $\sqrt{m^2}/(m^2 - 1)$ be added, *i.e.*,

$$p = \frac{E}{\sqrt{3}}\frac{t}{R}\sqrt{\frac{m^2}{m^2 - 1}}, \qquad \frac{1}{m} = \text{Poisson's ratio}$$

Southwell has also investigated (using the theory of thin shells) the conditions under which the lobed form of failure is produced and has given the formula

$$p = \frac{E}{\sqrt{3}}\frac{t}{R}\sqrt{\frac{m^2}{m^2 - 1}}\left(\frac{K^2 - 1}{K^2 + 1}\right),$$

where K is the number of lobes formed He has also given a curve representing the way in which the collapsing load depends on the length of the strut.

Dean has also obtained these formulæ by other methods of attack

In a later portion of this paper detailed consideration is given to these formulæ. It is suggested that an oversight in the interpretation by Southwell of his general analytical result has led him to omit consideration of the conditions which are most likely to occur in practice and that in consequence his formula for the lobed type of failure is incorrect. Dean appears to have fallen into the same error.

It appears that, correctly interpreted, Southwell's general equation yields the same value for the minimum collapsing load in the lobed as in the symmetrical type of failure.

From the theoretical investigations it is clear that only very thin tubes may be expected to fail by elastic instability. Taking mild steel, for example, for which the proportional elastic limit in compression may be 20 tons per square inch, the ratio of t/R which would be required to produce failure by elastic instability under this load per square inch would be about 1/400, *i.e.*, an 8-inch tube 1/100-inch thick. Such tubes are not met with in ordinary engineering work, so that the mathematical investigations would suggest that in tubes of ordinary proportions the proportional elastic limit of the material would be reached before any question of failure by elastic instability could arise. For stresses greater than the proportional elastic limit the formulæ cease to apply, since they are, of course, based upon the assumption of linearity between stress and strain. In some materials, however, the deviation from linearity of the stress-strain curve between the proportional elastic limit and the yield is so small that if the yield stress is used the formulæ will probably give a reasonable estimate of the value of t/R above which this type of deformation would not be expected to occur as a primary cause of failure.

It should be noted that these formulæ give a value for the load which will produce a condition of instability in the walls. Strictly interpreted, this is not necessarily a collapsing load, just as in the case of a solid strut the Euler load is not really a collapsing load. If the conditions could be properly chosen it would be possible to apply the Southwell load and note the formation of the bulge or lobes, and, on releasing the load, find the tube uninjured, *i.e.*, exactly as one may test a very long strut under its Euler load. Unfortunately, the difficulty of realising the conditions is so great that if any deflection does occur the material is probably strained beyond its proportional elastic limit and complete collapse generally occurs.

Experimental Investigations.

Lilly.—The most extensive series of experiments known to the author are those of Lilly, who used tubes of mild steel varying from ⅜-inch diameter 22 gauge to 1-inch diameter 18 gauge. From his results he deduced that the collapsing load was given by the formulæ

$$f = \mathrm{F}\Big/1 + \mathrm{K}\left(\frac{\rho}{t}\right)^2,$$

where f = load per square inch at collapse, F = strength in compression,

K = constant = 1/60 for mild steel, ρ = radius of gyration, t = thickness. In another paper he modified this to

$$f = \mathrm{F}\Big/1 + \mathrm{K}\left(\frac{\rho}{t}\right) \qquad (\mathrm{K} = 1/8 \text{ for mild steel}).$$

It is difficult to discuss these experiments since no stress-strain diagrams of the material in tension or compression are given and all the tubes are treated as if they were of the same material, despite the fact that the tensile strength varied from 62,000 lbs. per sq. inch to 80,000 per sq. inch.

The present author has discussed elsewhere* the importance of stress-strain diagrams in dealing with strut tests and the inadvisability of using specimens from different batches of material which have been subjected to different amounts of cold work when the object of the experiment is to ascertain how a material behaves under particular stress distribution.

Mason.—Probably the most carefully conducted tests bearing on the problem are those described by Mason in his paper on mild steel under combined stress. In the course of this research Mason carried out simple compression tests on tubes 2¾-inch bore 10 gauge (*i.e.*, $t/\mathrm{R} = 0 \cdot 089$) and tubes 3-inch bore 14 gauge (*i.e.*, $t/\mathrm{R} = 0 \cdot 052$). The yield in compression as obtained by an extensometer was designated Y_1 and as obtained by compasses designated Y_2. In some of the tests on the thinner tubes ($t/\mathrm{R} = 0 \cdot 052$) apparent yielding (attributed by the author to flexure, presumably the formation of a wrinkle) was observed at a stress below that at which the yield could be detected with a pair of compasses. In several other tests, however, on specimens cut from the same tube, the difference between Y_1 and Y_2 was "only about the same difference as between similar yield stresses of the tube in tension."

It may also be noted that in one series of tests on specimens from the same tube, Y_2 was 20·5 tons per square inch, Y_1 for a tube 19 inches long was 19·25 tons per square inch and for tubes 4 inches long 14·7 and 16·1 tons per square inch respectively. In another series also the longest tube gave a value of Y_1 much greater than the shorter tubes and greater than several of the values of Y_2 on other tubes. The length of the tube in every case was many times the wave-length of a wrinkle, so that the present author considers that the anomalies given above are probably due to unknown inequalities in the stress distribution which would be greater with a short specimen than with a long one, and that the actual wrinkle (if produced) was due to larger stresses than the average load per square inch recorded.

---

* 'Inst. Civ. Eng.,' Selected Paper 28.

For the other series of tubes $t/R = 0\cdot087$, in no case did any sign of flexure occur and the collapsing load was greater than the yield.

Barling and Webb.—These authors describe tests carried out at the Royal Aircraft Factory on annealed mild steel tubes. They defined the wrinkling stress as the stress at which the relation between the stress and the strain departs appreciably from the straight-line law, and to determine this point they used apparatus which gave autographic records of the relation between the load and the strain. They state that for $t/R =$ or $> 0\cdot06$ the wrinkling stress is constant at 30,000 lbs. per square inch, and for $t/R < 0\cdot06$ the wrinkling stress is directly proportional to t/R.

Popplewell and Carrington.—These authors carried out an investigation on high tensile steel tubes at the request of the Superintendent of the Royal Aircraft Factory. They adopted the definition of Messrs. Barling and Webb for the wrinkling stress and followed them also in using autographic diagrams, though they checked some of them with an extensometer. They tested the material in the hardened and annealed conditions.

For the hardened tubes the value of the wrinkling stress as defined above was uniform at about 50 tons per square inch between $t/R = 0\cdot049$ and $t/R = 0\cdot1$, and for values of $t/R > 0\cdot1$ the stress diminished by a linear law. For the annealed tubes the wrinkling stress was directly proportional to t/R up to $t/R = 0\cdot1$, and above this value was approximately uniform at 30 tons per square inch.

These experiments, and also those of Messrs. Barling and Webb, are open to the same criticism as Lilly's in that tubes of different sizes $\frac{3}{4}$-inch 22 gauge to 2-inch 20 gauge are treated as though they were identical material in spite of the fact that tensile tests showed that a hardened $1\frac{1}{4}$-inch 20 gauge tube had an ultimate stress of 104 tons per square inch, whilst a hardened $1\frac{1}{4}$-inch 12 gauge tube had an ultimate stress of only 75 tons per square inch. The inadvisability of this is also very clearly shown by the result Messrs. Popplewell and Carrington give of a test on a thin tube in the hardened condition which was turned down from a thick one. This thin tube ($t/R = 0\cdot055$) gave nearly the same value as the thick one ($t/R = 0\cdot17$), *i.e.*, about 36 tons per square inch as against about 39, whilst the other tubes of the same t/R as this thin one gave values of about 50 tons per square inch. It is thus clear that if all the tests had been made on tubes of the same material (*i.e.*, by turning down the thick ones) the results would have been the same over this range $t/R = 0\cdot055$ to $0\cdot17$, and there would have been no discontinuity at $t/R = 0\cdot1$. Moreover, the fact that over such a large range of t/R the value of what they call the wrinkling stress is constant is probably an indication of a stress failure rather than failure by elastic instability, since this must surely depend upon t/R.

From this review of the available experimental work it is clear that there is a very great diversity of opinion as to the value of the collapsing load for a tubular strut.

Author's Experiments.

The author's experiments have been made on a variety of different tubes, but the principal results can be stated by reference to four sets of experiments, *i.e.*,

A. Experiments on nickel-chrome steel tubes.
B. ,, ,, mild steel tubes.
C. ,, ,, large tubes made from thin steel strips.
D. ,, ,, tubes of silver spruce.

Experimental Methods.—In testing specimens in compression it is important to take precautions to ensure that the stress in the portion of the specimen under investigation is as uniform as possible. The ordinary compression plates of a testing machine are not necessarily parallel to one another at the beginning of a test, nor will they remain parallel under increased loading. In order to avoid inequalities of stress from this cause, the apparatus illustrated in fig. 1

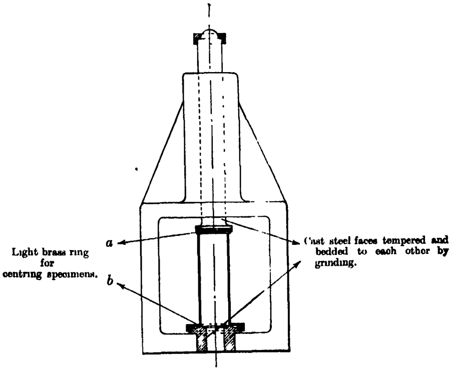

a
Light brass ring for centring specimens.
b
Cast steel faces tempered and bedded to each other by grinding.

FIG. 1.—Apparatus for maintaining Parallelism of Ends (× ½).

was made. The construction is clear from the diagram. Before testing, each specimen was ground against the faces *a* and *b*, so as to secure good initial bedding. A special extensometer was constructed for a gauge length of 1 inch. The construction is obvious from the diagrammatic sketch in fig. 2 and the photographs in figs. 3 and 4 (Plate 10). This apparatus was used on all the tubes about 1¼-inch diameter. For larger specimens no similar apparatus was available, so each specimen was provided with an end in which was centred a hard steel ball. Each compression plate was provided with a hard steel flat plate through which the load was transmitted (see fig. 5). If the tubes are uniform in themselves this secures a good approximation to uniform loading, but if the thickness is variable there will be a variation in the stress on a cross section, *i.e.*, the loading is virtually slightly eccentric.

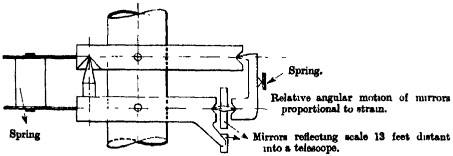

Spring.
Relative angular motion of mirrors proportional to strain.
Spring
Mirrors reflecting scale 13 feet distant into a telescope.

FIG. 2.—Diagrammatic Sketch of Extensometer.

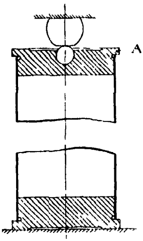


FIG. 5.

Nickel-Chrome Steel Tubes.

Preliminary tests on 1¼-inch diameter 18-gauge tubes of nickel-chrome steel indicated that for tubes having $t/R = 0{\cdot}083$ the hardened tubes showed no signs of collapse under a stress of 72 tons per square inch, and that the annealed tubes had a yield which varied from 28 to 39 tons per square inch. The stress-strain diagrams of specimens from different tubes when tested as received from the makers indicated very considerable differences which suggested improper heat treatment. An investigation was therefore made into the heat treatment of this material. The investigation showed the importance of three conditions, *i.e.*,

(*a*) Annealed.*
(*b*) Hardened from 800° C.
(*c*) Hardened from 800° C. and then tempered at 400° C.

Three series of tests were, therefore, carried out on tubes with varying values of t/R. In every case the specimens in each series were from the same tube and the various ratios of t/R were obtained by machining or grinding the tube.

The stress-strain diagrams and the collapsing loads plotted against t/R are given in fig. 6 and photograph of typical collapsed specimens in fig. 7 (Plate 11).

In the case of the thinnest tubes, which were only 0·005 inch thick, the collapsing load only was obtained.

For the annealed specimens the proportional elastic limit varied from 35 to 40 tons per square inch, and the yield about 45 tons per square inch. For the hardened specimens the proportional elastic limit is low, *i.e.*, from 28 to 43 tons, but there is no yield, the stress-strain curve resembling in shape that of cast iron. For the tempered specimens the proportional elastic limit is high, 58 and 78 tons per square inch, and at a stress of 80 to 90 tons per square inch the stress-strain curve has a very considerably increased slope.

The relation of the collapsing loads to t/R may most easily be summarised by reference to the theoretical formulæ for the collapse by wrinkling, *i.e.*, $p = \frac{\mathrm{E}}{\sqrt{3}}\frac{t}{\mathrm{R}}\sqrt{\frac{m^2}{m^2-1}}$, which is also plotted in fig. 6. For the tempered and annealed specimens the experimental values are about 0·4 of the theoretical ones up to such values of t/R as make this value (*i.e.*, 0·4 p) slightly greater than the proportional elastic limit. For greater values of t/R the collapsing load per square inch is nearly uniform at about 92 tons per square inch up to $t/\mathrm{R} = 0{\cdot}08$ for the tempered specimens, and rises slightly with t/R up to 50 tons at $t/\mathrm{R} = 0{\cdot}08$ for the annealed specimens.

This generalisation would also cover the mild steel tubes, as is shown by the results of a series of machined specimens also plotted in this diagram.

The results on the hardened tubes are more irregular, probably the low result

---

* The actual treatment was hardening from 800° C. and subsequent tempering at 650° C., and would be more correctly described as "softened." The term annealed has been retained as it is used in other papers for this condition.

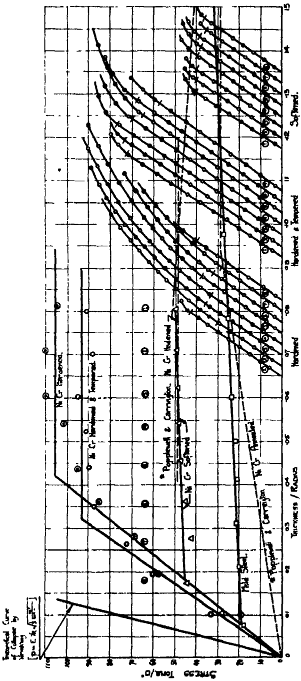


FIG. 6.—Compression Tests on Nickel-Chrome Steel Tubes.*

for $t/R = 0 \cdot 01$ is due to unequal stress distribution. Apart from this one, however, the results are not greatly different from those of the tempered specimens up to $t/R = 0 \cdot 04$. For larger values of t/R the collapsing load is about 105 tons per square inch. In considering the results it must be remembered that owing to the enlarged ends there is an unknown additional stress produced which is not included in the theoretical analysis

---

* These tests are for the first departure from linearity and should be compared with the stress-strain diagrams.

These results may be compared with those of Messrs Popplewell and Carrington, which were made on tubes to the same specification, and supplied, as were the author's tubes, by the Superintendent of the Royal Aircraft Factory. The results of Messrs. Popplewell and Carrington are shown by dotted lines in fig. 6. Their results refer to the first deviation from the straight line.

If these results be compared with the proportional elastic limits as found by the author, it will be seen that there is considerable divergence. The author considers that Messrs. Popplewell and Carrington's results are too low for the annealed tube and too high for the hardened tube, and that this is due to their method of test not securing a good distribution of stress and to their autographic diagram and extensometer being less sensitive than the author's. Moreover, they used tubes of different diameters and thicknesses and assumed that they were dealing with the same material. It has previously been pointed out that their published results show that if they had machined the specimens from the same tube they would have had no marked break at $t/R = 0 \cdot 1$ in the case of the hardened tubes The author, therefore, dissents entirely from the conclusions which Messrs. Popplewell and Carrington drew from their own experiments.

For almost identical reasons the experiments of Messrs. Barling and Webb, which were carried out on similar lines to those of Messrs. Popplewell and Carrington, and which Messrs. Popplewell and Carrington claim are verified by their investigations, are also misleading.

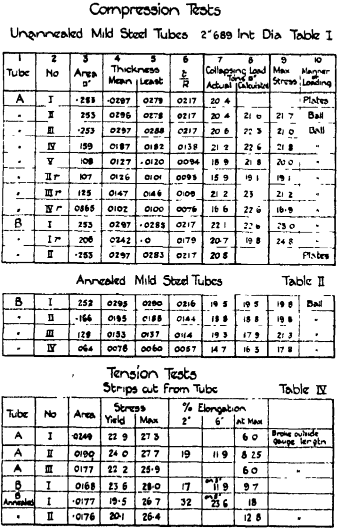
Mild Steel Tubes, 2¾-inch Diameter.

For these experiments the method of loading was that shown in fig. 5. This method avoids serious inequalities of strain at the ends, and though the strain is not uniform unless the tube is of uniform thickness, an estimate of the variation can be made if the variation in the thickness is known

For all these tests autographic records of load and strain were obtained. The actual decrease in length of the tube under test was multiplied by a system of levers so as to give on the diagram a magnification of about 8.

In taking the diagram one of the levers used to magnify the motion was kept in contact with the underside of the flange A. On account of the variation in thickness the tube rarely collapsed uniformly with the ends parallel, so the motion of the end of the lever does not give a correct measure of the mean longitudinal motion after yield has occurred. In a few cases the lever was either pushed away by the formation of a fold or the end tilted so much in collapsing that the diagram was reduced to the single line traced during the loading period.

The specimens were about 6 inches long and several of them were machined on the outside to give various ratios of t/R, the lowest being 0·0057. The unannealed specimens IV and V were machined over their entire length.



Compression Tests

Unannealed Mild Steel Tubes 2″689 Int Dia Table I

1 Tube	2 No	3 Area □″	4 Thickness Mean	5 Least	6 $\frac{t}{R}$	7 Collapsing Load Tons □″ Actual	8 Calculated	9 Max Stress	10 Manner of Loading
A	I	·253	·0297	0279	0217	20 4			Plates
"	II	253	0296	0278	0217	20 4	21 6	21 7	Ball
"	III	·253	0297	0288	0217	20 6	22 3	21 0	Ball
"	IV	159	0187	0182	0138	21 2	22 6	21 8	"
"	V	108	0127	·0120	0094	18 9	21 8	20 0	"
"	II r	107	0126	0101	0093	15 9	19 1	19 1	"
"	III r	125	0147	0146	0109	21 2	23	21 2	"
"	IV r	0865	0102	0100	0076	16 6	22 6	16·9	"
B	I	253	0297	·0285	0217	22 1	22 6	23 0	"
"	I r	206	0242	·0	0179	20·7	19 8	24 8	"
"	II	·253	0297	0283	0217	20 8			Plates

Annealed Mild Steel Tubes Table II

B	I	252	0295	0290	0216	19 5	19 5	19 8	Ball
"	II	·166	0195	0188	0144	18 8	18 8	19 8	"
"	III	129	0153	0137	0114	19 3	17 9	21 3	"
"	IV	064	0076	0060	0057	14 7	16 3	17 8	"

Tension Tests

Strips cut from Tube Table IV

Tube	No	Area	Stress Yield	 Max	% Elongation 2″	 6″	 At Max	
A	I	·0249	22 9	27 3			6 0	Broke outside gauge length
A	II	0190	24 0	27 7	19	11 9	8 25	
A	III	0177	22 2	25·9			6 0	"
B	I	0168	23 6	28·0	17	on 8″ 11 9	9 7	
B Annealed	I	·0177	19·5	26 7	32	on 8″ 23 6	18	
"	II	·0176	20·1	26·4			12 8	"

FIG. 8.—Compression and Tension Tests—Tables.

After the first test the collapsed portions (which were at the ends) of Nos. II III, IV were cut off, and the remainder machined down over the middle 4 inches. These tests were denoted by II_r, III_r, IV_r. When necessary, the tubes were cut in order to determine the variation of thickness in the region close to that which had collapsed.

The results of these tests are given in fig. 8, some of the autographic diagrams in fig. 9, and photographs of some of the collapsed specimens in

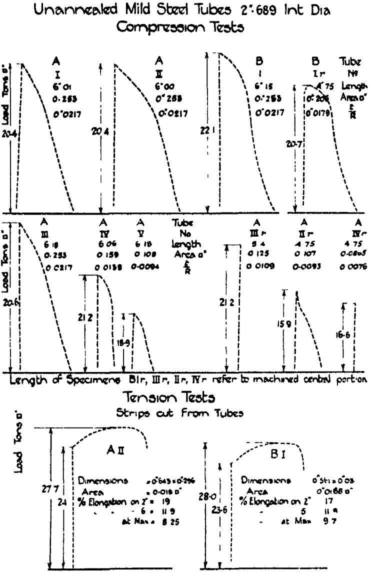


FIG. 9.—Compression and Tension Tests—Autographic Diagrams.

figs. 10 and 11 (Plate 11). The results of some tension tests made on a strip cut from the tube are also given in fig. 9.

In column 7 is given the actual collapsing load per square inch, and in column 9 the calculated maximum stress, allowing for the eccentricity of loading produced by the variation of thickness. The tubes were too thin to obtain the real yield in compression, but if we assume it to be the same as that observed in the tension test and then calculate the load per square inch required to produce this stress in the specimens, we obtain the figure given in column 8.

It will be seen from 7 and 9 that, although the load per square inch at collapse varied from 16 to 21 tons per square inch for unannealed tubes and from 15 to 20 tons for annealed tubes, the maximum stress is but little below the yield in tension. It may also be noted that this yield in the unannealed tube is not a very well-marked one, and the determinations may easily be on the high side.

It should be noted that the specimens generally collapsed near the ends or where the section was changed, for at this place there is an additional stress introduced on account of the material being prevented from expanding freely (by reason of friction at the ends, or change of section), as it is permitted to do at points in the middle of the specimen The case of the annealed specimen No. II is interesting in that collapse actually occurred near the end where the section was 0·188 square inch, and the stress 16·6 tons per square inch, though the tube was turned down to give an area of 0·166 on the centre 4½ inches. This central portion under a stress of 18·8 tons per square inch showed no signs of collapsing.

The photograph of the annealed specimen No. I (fig. 11) is interesting, for it shows several Luder's lines extending over a length greater than the wave-length of a fold, thus showing that these portions of the tube were yielding, though no sign of wrinkling at these places could be detected.

It is of interest to note the shape of the collapsed portions of the tube. In those places where failure occurred at the ends, the plug prevented any deformation other than a simple fold outwards, as shown in the annealed specimens I and II, fig. 11; where, however, the tube was free to take up the shape conditioned by the stress-strain relations, it collapsed by caving in at regular intervals round the circumference.

For tubes having :—

t/R = 0·022 the pitch of the indentation was ¼ circum.
= 0·01 and 0·017 the pitch of the indentation was ⅙ circum.
= 0·0076 the pitch of the indentation was ⅛ circum.
= 0·0057 the pitch of the indentation was $\frac{1}{9}$ circum.

It may be noted that thick tubes of mild steel give an autographic diagram very similar to an autographic diagram in tension. This is shown by the diagram in fig. 12 (Plate 12). After yield has been passed it can generally be noticed that the walls of the tube have bulged, but complete collapse does not occur till the load is much greater than the yield. Generally a specimen which gives a diagram showing a horizontal line at yield collapses with a characteristic uniform bulge, whilst one giving only a peak will cave in in several places. This is shown in fig. 14 (Plate 12), which represents a remarkably uniform failure.

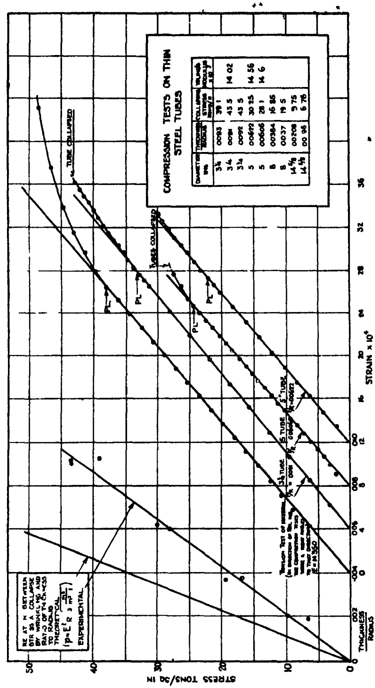
Tubes made from High Tensile Steel Strip.

In order to obtain tubes which might be expected to fail by elastic instability rather than by reason of the stress distribution, a series of tubes was made from steel strip such as is used for metal construction of aircraft. The particular steel used, for which the author is indebted to Major H. Wylie, has a limit of proportionality in tension of about 40 tons per square inch and is rolled to about 0·015 inch thick. The tubes were made by wrapping the strip (which was about 5 inches wide) on a mandril so that the edges butted and then soldering the joint. The ends were faced up whilst on the mandril so as to secure a good bedding on the end-plates, which were similar in design to those employed for the 2¾-inch tubes described above.

It is very difficult to get a good compression test of the material when in the form of such thin strips. An attempt was made to make a test piece by soldering together a large number of pieces and then machining the resulting block to a square section. The result, however, was not satisfactory.

On the 3-inch and 5-inch specimens an extensometer was used, and the stress-strain curves are given in fig. 15. The tension stress-strain curve for the material in the direction of the length of the strip is also given. It should be noted, however, that this test is at right angles to the direction in which the stress is applied in the compression tests. Tests were made in tension across the strip, but the difficulties of securing a reasonably accurate stress distribution on such short specimens gave results which were considered fictitious. It is, therefore, impossible to say what is the real proportional elastic limit in compression in the direction in which the stress was applied.

The results are all given in fig. 15. It will be seen that the results are all about 0·6 of those given by Southwell's formula for a bulge type of collapse and that the 3-inch tubes $t/R = 0{\cdot}0093$ sustained a stress of 43·5 tons per square inch, which is probably above the limit of proportionality of the material.



COMPRESSION TESTS ON THIN STEEL TUBES

DIAMETER INS.	THICKNESS / RADIUS	COLLAPSING STRESS TONS/□"	YOUNG'S MODULUS × 10³
3¼	0093	39 1	
3¼	0091	43 5	14 02
3¼	0092	43 5	
5	00622	30 25	14 56
5	00606	28 1	14 6
8	00384	16 85	
8	0037	19 5	
14⅝	00208	9 75	
14⅝	00 96	6 76	

FIG 15

It has to be remembered that the strip is not exactly straight, and that the tubes cannot be exactly circular. These two defects would, of course, lead to values less than the theoretical ones.

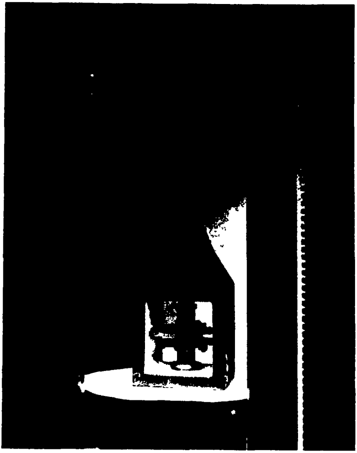


Fig. 3

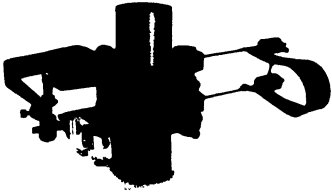


Fig. 4.

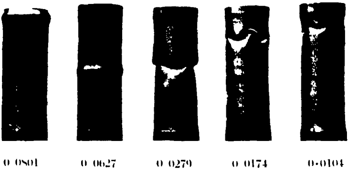


Fig. 7.—Values of Thickness/Radius. Ni-Cr Steel. Unannealed.

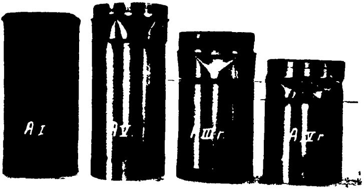


Fig. 10.—Unannealed Mild Steel Tubes.

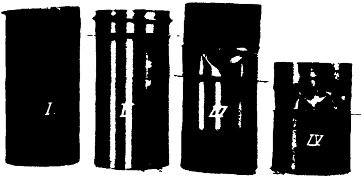


Fig. 11.—Annealed Mild Steel Tubes.

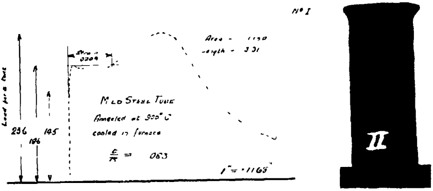


FIG. 12

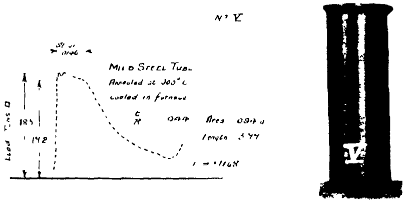


FIG. 13.

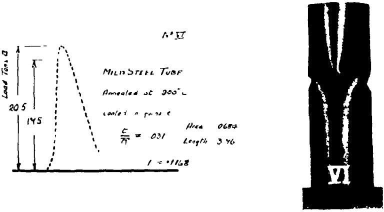


FIG. 14

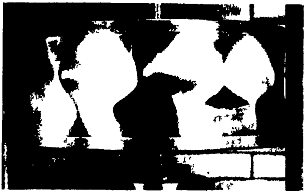


FIG. 16.—Photograph of Large Tube, before removing Collapsing Load.

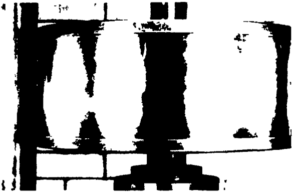


FIG. 17.—Photograph of Large Tube, after removing Collapsing Load.

0·187 0·082 0·032

FIG. 21.—Values of Thickness/Radius. Tubular Struts—Spruce.

The largest tubes were tested in a special machine in series with a specimen of tool steel, to which an extensometer was attached. This enabled the deformation at failure to be restricted to a smaller value than it would have had if tested in the usual testing machine. A photograph, fig. 16, was taken of the tube just after failure, and another, fig. 17, when the load had been relieved (Plate 13). These show that a considerable amount of the deformation was elastic, since the characteristic lobes have nearly disappeared on the removal of the load.

Comparison of Experimental Results with the Theoretical Formula.

This series of tests is the only one that really comes within the scope of the available theoretical analysis, and in particular of Southwell's theory of the formation of lobes obtained by using the theory of thin shells. The author has recently had occasion to examine Southwell's work in the light of the experimental results just described, and has found reasons for considering that an oversight in the interpretation of a long equation (expressing the results of the analytical investigations) has rendered Southwell's formula* (105) and diagram, fig. 6 (p. 235), inapplicable to the problems here considered. It will be necessary to consider this point at some length.

Southwell obtained the equations which follow, and these are supported by an independent investigation recently published by W. R. Dean

$$\psi = -\frac{m^2-1}{m^2}\frac{p}{\mathrm{E}} = -\frac{m^2-1}{m^2}\frac{\mathrm{P}}{4\pi a t_1 \mathrm{E}}, \tag{96}$$

$$\begin{aligned}&\psi^2\left[\tfrac{1}{2}(m+1)\,\mathrm{K}^2q^4/m\right]+\psi\left[\tfrac{1}{2}(m-1)\,q^2\{(\mathrm{K}^2+q^2)^2+\mathrm{K}^2\}/m\right]\\&+\tfrac{1}{2}(m^2-1)(m-1)\,q^4/m^2+\tfrac{1}{3}t_1^2/a^2\{\psi q^2[\tfrac{1}{2}(m+1)\,\mathrm{K}^2\{(\mathrm{K}^2+q^2)^2-\mathrm{K}^2\}/m\\&+\tfrac{1}{2}(m^2-7m+4)\,\mathrm{K}^2q^2/m^2+(m-1)\,q^4/m]\\&+\tfrac{1}{2}(m-1)/m\,[(\mathrm{K}^2+q^2)^4-q^6/m-\mathrm{K}^2\{2\mathrm{K}^4+7\mathrm{K}^2q^2+(7m^2+m-2)\,q^4/m^2\}\\&+\mathrm{K}^4+3\mathrm{K}^2q^2+(2m^2-1)q^4/m^2]\}=0,\end{aligned} \tag{98}$$

where p = load per square inch, E = Young's modulus, $1/m$ = Poisson's ratio, K = number of lobes, $q = 2\pi a/\lambda$, λ = wave-length of a lobe, a = radius = R, $t_1 = \frac{1}{2}$ thickness of tube, t = thickness of tube.

The cases to be considered are (1) K = 0, (2) K = 1, 2, 3, etc.

---

[*] The references in this section of the paper are to Southwell's original paper, 'Phil. Trans.,' A, vol. 213, and his nomenclature has been adopted except that t_1 is used for the half thickness of the tube instead of t.

When $K = 0$ the collapsing load is the same as that obtained by his more rigorous analysis in the early part of his paper, *i.e.*,

$$p = E\frac{2t_1}{a}\sqrt{\frac{1}{3}\frac{m^2}{m^2-1}} = E\frac{t}{R}\sqrt{\frac{1}{3}\frac{m^2}{m^2-1}}.$$

This formula is obtained from equation 98 in virtue of simplifications which can be introduced on account of the fact that p/E must be very small in any case of practical importance. When $K = 0$, q must be great in order that this condition may be satisfied, and only the highest powers of q need be retained. When $K \neq 0$ Southwell concluded (p. 234) that q must be small (*i.e.*, the wave-length must be great), and his formula (105) and fig. 6 (p. 235) were obtained on this assumption. But although a small value of q (when t_1/a is very small) may correspond with a possible value of p/E, it is not a necessary condition for this result. A small value of p/E can also occur with a large value of q (small wave-length), and this is, in fact, the dominating condition in the problem. We therefore require to examine the consequence of (98) when the conditions are not limited in the way Southwell considered.

Examination of the relative value of the terms in any practicable case shows that the equation may to a close approximation be reduced to

$$\frac{p}{E} = \left(\frac{q}{K^2+q^2}\right)^2 + \frac{1}{3}\left(\frac{t_1}{a}\right)^2\frac{m^2}{m^2-1}\left(\frac{K^2+q^2}{q}\right)^2 \tag{1}$$

$$= \frac{1}{\alpha^2} + A\alpha^2, \tag{2}$$

where $\alpha = (K^2+q^2)/q$, $A = \frac{1}{3}t_1^2m^2/a^2(m^2-1)$.

We desire to know the minimum value of p/E since this will be the load at which instability will occur, and the corresponding values of K and q. For a minimum

$$\frac{d(p/E)}{dq} = \left(2A\alpha - \frac{2}{\alpha^3}\right)\left(1 - \frac{K^2}{q^2}\right) = 0, \tag{3}$$

and

$$\frac{d(p/E)}{dK} = \left(2A\alpha - \frac{2}{\alpha^3}\right)\left(\frac{2K}{q}\right) = 0. \tag{4}$$

To satisfy both (3) and (4), $\alpha^4 = 1/A$. (5)

Substituting in (2)

$$\frac{p}{E} = \sqrt{A} + \frac{A}{\sqrt{A}} = 2\sqrt{A} = 2\frac{t_1}{a}\sqrt{\frac{1}{3}\frac{m^2}{m^2-1}},$$

i.e., exactly the same equation as that for case 1 when $K = 0$.

From (5) we can find the value of q (for any given value of K) when the collapsing stress is a minimum.

$$(K^2 + q^2)/q = \sqrt[4]{(1/A)}. \tag{6}$$

$$q = \{ \sqrt[4]{(1/A)} \pm \sqrt{(1/A)^{\frac{1}{2}} - 4K^2}\}/2. \tag{7}$$

For real values of q

$$K^2 < 1/(4\sqrt{A}). \tag{8}$$

From (6), (7) and (8) it is evident that the tube can fail in any number of lobes less than the maximum given by (8), and that for any given value of K there are two possible values of q.

This is clearly shown in fig. 18, in which the collapsing stress for a particular tube is plotted against $1/q$. The maximum value of K, to give a minimum collapsing stress, for this particular tube, is 20, so that it could fail by the formation of 20, or any smaller number of lobes. Curves for K = 20, K = 15, and K = 10 are shown plotted, and it will be seen that each curve has two minima for two different values of q, and that all these minima occur at the same collapsing stress. The two minimum values for K = 20 are practically coincident, and the second value for K = 10 is to the right of the diagram. The curve for K = 30 is also shown, and it will be seen that its minimum is considerably higher than the minima of all curves in which K is less than 20.

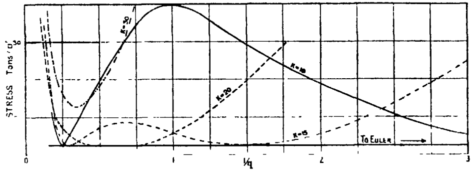
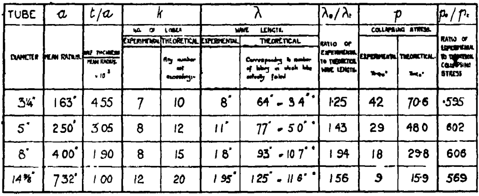


FIG. 18.—Collapsing Stress of Thin-Walled Steel Tube. $2t/a = 0{\cdot}002$, $E = 3 \times 10^7$, $m = 10/3$.

A comparison of the experimental results with the theoretical ones is set out in fig. 19. It will be seen that the actual number of lobes into which the tube collapsed was always less than their minimum, and that the wave-length as measured after collapse had occurred was always greater than the theoretical value. The actual collapsing load was about 0·6 the theoretical. It must, however, be remembered that these tubes deviate from the assumption in two important particulars, for they cannot be considered as either truly circular or quite straight. Moreover, there is also the complication previously pointed out due to the friction at the ends. All these effects would tend to reduce the collapsing load below that given by the theoretical analysis. It would be an expensive job to machine tubes of these sizes 14 inches diameter 0·014 inch thick from the solid, and the author has not attempted it. The results obtained, *i.e.*, 60 per cent. of the theoretical, are regarded as a substantial step towards the verification of the Southwell formula.



TUBE	a	t/a	k		λ		λ_e/λ_t	p		p_e/p_t
			No. of lobes		Wave length			Collapsing stress.		
			Experimental	Theoretical	Experimental	Theoretical				
Diameter	Mean radius.	Half thickness / mean radius × 10³		Any number not exceeding—		Corresponding to number of lobes on which tube actually failed	Ratio of experimental to theoretical wave length	Experimental Tons/in.²	Theoretical Tons/in.²	Ratio of experimental to theoretical collapsing stress
3¼″	1 63″	4 55	7	10	8″	64″ - 3 4″ *	1·25	42	70·6	·595
5″	2 50″	3 05	8	12	11″	77″ - 5 0″ *	1 43	29	48 0	602
8″	4 00″	1 90	8	15	1 8″	93″ - 10 7″ *	1 94	18	29·8	606
14⅝″	7 32″	1 00	12	20	1 95″	125″ - 11 6″ *	1 56	9	15·9	569

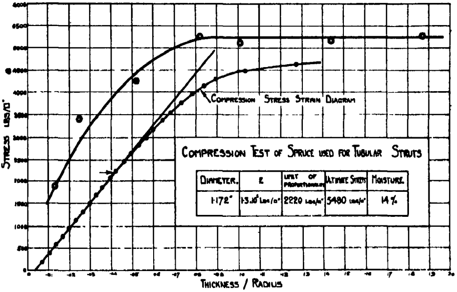
FIG. 19.—Comparison of Theoretical and Experimental Results.

* Large value of λ only approximate. Equation not accurate when q is small ($q = 2\pi a/\lambda$).

Timber.—One series of tests on tubes made from spruce was carried out. Thin strips from the same plank were carefully planed to a fairly uniform thickness, about 0·025 inch. The various tubes were made by wrapping the necessary number of strips round a mandril having first spread a fine coating of glue on all the faces that were to come together. To minimise the risk of failure at the ends an extra strip, ½ inch wide, was glued on. Care was taken that the grain of the timber was parallel to the axis of the mandril.

On a solid specimen from the same plank a compression test was carried out. The results are given in fig. 20 and a photograph of some of the collapsed specimens in fig. 21 (Plate 13).

It will be seen that the collapsing stress is uniform and practically that of the solid specimen for all values of $t/R > 0\cdot08$. It should be noted that in the case of timber, which has such different elastic properties in different directions, there is as yet no theory giving the collapsing load in terms of the elastic constants. In view, however, of the extensive use of timber in aeronautical work, it was considered advisable to carry out this series so as to determine the probable limit of t/R for hollow struts.



COMPRESSION TEST OF SPRUCE USED FOR TUBULAR STRUTS

DIAMETER.	E	LIMIT OF PROPORTIONALITY	ULTIMATE STRESS	MOISTURE
1·172"	1·3.10⁶ lbs/□"	2220 lbs/□"	5480 lbs/□"	14 %

FIG. 20.—Spruce Tubular Struts.

Strut Tests.

Several series of tests of free-ended struts were carried out on mild steel and nickel-chrome steel tubes. For each specimen the tubes were cut into lengths, the ends faced up and provided with a plug, in which a steel ball was centred. During the tests the load on the strut was transmitted through hardened flat plates (usually flats ground on 1-inch balls), as shown in fig. 22. The length of the strut was taken as the distance between the centres of the balls. The load was therefore applied at the end of the strut in the centre line of the bore. Since the tubes were not uniformly thick, this introduced a small eccentricity.

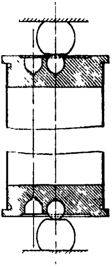


FIG. 22.

It will be convenient to treat long and short struts separately, using the term "long" to apply to struts of such a length that the Eulerian value is much less than the elastic limit of the material. This distinction is convenient since the factors producing failure are different in the two cases. In a long strut failure is chiefly determined by the elastic instability which occurs at the Euler load, whilst in the short strut failure is determined by stress considerations.

Long Struts.—The critical load for a long strut is that given by Euler's formula. Three tests were made, and the results are given in Table I.

Table I.

Materials.	L/K	E tons per sq. in.	Elastic limit, tons per sq. in.	Yield stress, tons per sq. in.	Collapsing load, tons per sq. in.		Percentage error compared with Euler.
					Actual	Euler	
I.—Mild steel	256	12,760	17·6 11·65 9·5	24	1·88	1·90	−1
II.—Hardened high tensile steel	119	12,400	16·3	None	8·49	8·60	−1·3
III.—Annealed high tensile steel	119	13,330	33·0	38·5	8·61	9·25	−6·8

It will be seen that the experimental values agree very closely with the theoretical despite the inevitable eccentricity of loading to which the tubes are subjected owing to unequal wall thickness.

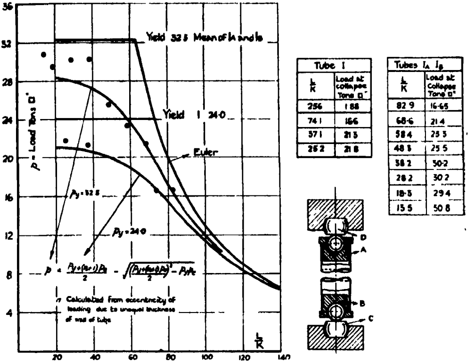
Short Struts.—It is convenient to consider materials having considerable ductility (such as mild steel) separately from those like hardened nickel-chrome steel, in which the ductility is relatively very small.

Short struts fail by reason of stress conditions. The author has shown[*] that for material having a drop of stress at yield, failure occurs when the stress at any point reaches the yield. The usual formulæ are only applicable for stresses within the elastic limit, but, in the case of mild steel, they may, without serious error, be used for stresses up to the yield. In all tests of tubes there is likely to be eccentricity of loading due to the unequal wall thickness as well as that due to imperfect end fittings. To represent the results, therefore, either the usual eccentricity formula may be employed or the curvature formula, using Prof. Perry's approximation that a strut having an eccentricity of e is approximately equal to a strut bent initially into a sine curve and having a deflection at the centre as measured from the line joining the ends of $C_0 = \frac{9}{8}e$.

'Inst. Civ. Eng.,' Selected Paper 28.

Ductile Materials. Tubes 1¼ inch, 18 Gauge.

Mild Steel.—The results of two series of tests are given in fig. 23. The test pieces for each series were cut from a single tube and the yield in compression was determined by an extensometer test. If the specimens had been straight and axially loaded the results would have followed fairly closely the yield line and the Euler curve. Owing to the variation in thickness of the walls, however, the test pieces are loaded eccentrically by an amount approximately 0·027 inch. Assuming that failure will occur when the stress reaches the yield stress in compression of the material the curves for this eccentricity have been plotted and the results are seen to agree fairly closely with these curves.



Tube I	
$\frac{L}{K}$	Load at collapse Tons □"
256	1 88
74 1	166
57 1	21 5
25 2	21 8

Tubes I_A I_B	
$\frac{L}{K}$	Load at Collapse Tons □"
82 9	16·65
68·6	21 4
58 4	25 3
48 3	25 5
38 2	50·2
28 2	30 2
18·5	29 4
15 5	50 8

FIG. 23.—Mild Steel Tubes, 1¼ inch 0.0.18 gauge.

Annealed High Tensile Steel.—The results of a series of tests are given in fig. 24, series 1. It will be seen that, allowing for the inevitable eccentricity due to unequal wall thickness, the results agree very well with the eccentricity formula in which the stress term is the yield stress.

On the same diagram are given the results of several series under eccentric loading of various amounts. The experimental results are all above those given by the formula, very appreciably in cases of the short struts and slightly in cases of longer ones. This arises from the fact that when loaded with considerable eccentricity a large portion of the stress is due to bending, and

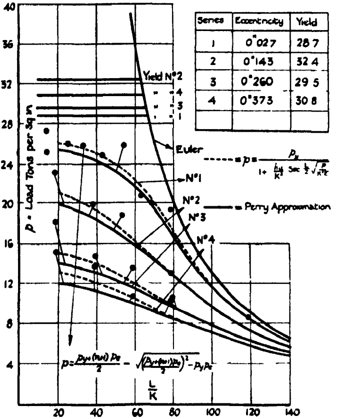


FIG. 24.—Eccentrically Loaded Struts—Annealed NiCr Steel, tubes 1½ inches out. diam., ·18 gauge.

hence with increasing eccentricity the behaviour of the strut tends to resemble that of a beam. Moreover, there is little drop of stress at yield in this material, and, therefore, the collapsing load is appreciably greater than that calculated on a basis of complete collapse when the yield stress is attained at any point. This point has been discussed by the author in his paper on struts.

Strut Tests on Mild Steel Tubes, 2½ inches O.D., 22 Gauge, t/R = 0·022.

The struts were provided with the same type of end as was used in the other series, by which the load was applied through the centre of the bore. The wall thickness was measured and the eccentricity of loading calculated.

The results are given in Tables II, III, and photographs of the portions of

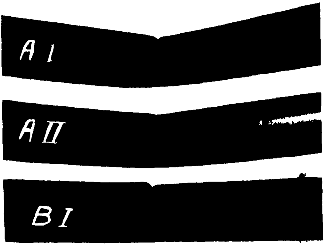


Fig. 25

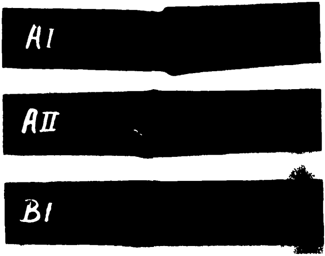


Fig. 26

(Facing p. 581)

the struts affected by the collapse are given in figs 25 and 26 (Plate 14) one view being taken in the plane of bending the other at right angles to it.

Table II Strut Tests

Tube	No	Area	Thickness		Eccentricity	l/K	Collapsing load		Max stress at collapse †
			Mean	Least			Actual (8)	Perry * (9)	(10)
A	I	0 253	0 0296	0 0280	0 037	33 8	20 7	21 2	22 2
A	II	0 253	0 0296	0 0284	0 625	33 8	10 9	10 3	22 4
B	I	0 253	0 0296	0 0277	0 044	94 2	12 7	13 0	21 5

* On assumption that the yield in tension the yield in compression

† Max stress = Load per square inch × $\left\{1 \quad \frac{ah}{K^2}\operatorname{Sec}\frac{\pi}{2}\sqrt{\left(\frac{p}{p}\right)}\right\}$

p = Load per sq in p = Euler an value h Eccentricity a Radius

Table III —Tension Tests (Strips out from Tube)

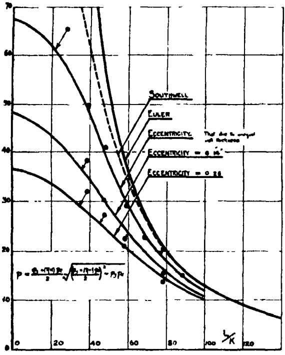
Tube	No	Area	Stress		Percentage elongation			—
			Yield	Max	2	6	At max	
A	I	0 0249	22 9	27 3			6 0	Broke outside gauge length
A	II	0 0190	24 0	27 7	19	11 9	8 25	
A	III	0 0177	22 2	25 9			6 0	
B	I	0 0168	23 6	28 0	17	On 5″ 11 9	9 7	
B Annealed	I	0 0177	19 5	26 7	32	On 5″ 23 6	18	
	II	0 0176	20 1	26 4			12 8	

The probable collapsing load has been calculated from the Perry formula, the yield in compression being taken as equal to the yield in tension From the actual collapsing load the maximum stress produced on the concave side has been calculated

It will be seen from an inspection of columns 8, 9 and 10 of Table II that in every case (i e , L/K = 33 8 33 8 94 2, and eccentricities 0 037 inch, 0 625 inch and 0 044 inch, respectively) the maximum stress is practically the yield stress, and that the Perry formula gives a good approximation of the collapsing load even with an eccentricity as great as 0 625 inch

Hardened High Tensile Steel.

The results of the test are given in fig. 27. It will be seen that for the series in which the eccentricity is only that due to unequal wall thickness the results follow the Southwell curve for axial loading, but are decidedly below it.



Eccentricity	L/K.	Collapsing load, tons/sq. in.
That due to unequal wall thickness	88·6	15 0
	78·2	20·5
	68·4	22·8
	59·0	29·2
	48·6	41·0
	39·6	49·6
	27·7	65·3
0·14"	78·0	15·6
	58·5	22·6
	47·4	30·3
	38·9	38·4
0·26"	78·1	13·7
	58·1	21·0
	47·8	27 3
	38 6	32·1

FIG. 27.—Hardened High Tensile Steel Tubes. 1¼ inches out. diam., 18 gauge.

There is no theoretical analysis available to enable the effect of eccentricity to be calculated for a strut made of a material of this type. For materials with very considerably greater ductility than that of this material in this condition, the author has obtained a fair agreement with experimental results by taking the stress term in the curvature formula as that at which the tangent to the stress-strain curve is one-third the value it has in the elastic portion of the curve. The method is artificial, but even in this case gives an approximate representation of the probable collapsing load, as is shown by the curves given in the diagram.

In the absence of a mathematical analysis in a material of this type it is not considered worth while to obtain an empirical formula to fit the results.

Conclusions.

From the work described in this paper the following conclusions may be drawn:—

I. For tubular struts of mild steel having a yield in compression of 22 tons per square inch the strength depends on the yield stress and not on the wrinkling stress provided $t/R \geqslant 0 \cdot 022$.

II. For short specimens of thin tubes tested in compression failure will occur either at the yield or at a stress which is some fraction of the Southwell value (p_s) for the load required to produce elastic instability in the walls. For the tubes tested the fraction varied from 0·4 for solid-drawn small tubes of mild steel and nickel-chrome steel to 0·6 for tubes made from the steel strip used in metal construction of aircraft. The value of this fraction doubtless depends upon the accuracy of the tubes, and even with perfect tubes would be less than unity on account of the additional stresses introduced in applying the load. In all cases of thin tubes the wave-length of a deformation is small, being considerably less than the radius of the tube.

III. For tubes having t/R greater than the value as determined by II above, the proportional elastic limit and also yield, where this is well defined, precedes wrinkling, and the particular type of deformation produced depends upon the ratio t/R.

From these conclusions the author considers that a reasonable estimate of the strength of a steel tubular strut can be made by using the eccentricity formula (Prof. Smith) or the modification by Prof. Perry and inserting as the stress term the lower of two values, *i e.*, the yield or an appropriate fraction of the Southwell value for collapse by elastic instability. In order, therefore, to avoid elastic instability as a primary cause of failure the value of t/R for steel tubes should be greater than $0 \cdot 4\ p_e$, *i.e.*, 0·006 for mild steel (20 tons yield). It is hardly likely that tubes of such a character will be used in ordinary engineering work, but this investigation has an important bearing upon the design of metal construction for aircraft. In this work one of the important elements of design is to avoid failure by elastic instability when compression stresses have to be taken. In order to do this flat surfaces are avoided and the flanges of girders, etc., are made curved or indented.

These conclusions derive considerable support from certain experiments carried out by Fairbairn* in 1846 in connection with the Conway tubular bridge.

---

* W. Fairbairn, 'Conway and Menai Tubular Bridges,' London (1849); E. Clark, Britannia and Conway Tubular Bridges,' London (1850).

He tested under a central load tubes built up from iron plates. The tubes varied from 12 inches diameter 0·048 inch thick to 24 inches diameter 0·095 inch thick, and the length between the supports was varied between 15 feet 6 inches and 31 feet. In Table IV are given the results of these tests, the experiments being arranged in order of the value of t/R. As the tubes were of wrought iron with an elastic limit of probably 14 tons per square inch, the tubes would require to have a t/R of 0·0018 (calculated from Southwell's theoretical formula) in order to fail from elastic instability. One would, therefore, on the basis of the author's work, expect them to stand up to a stress equal to the proportional elastic limit before failure by wrinkling. In only two cases did failure occur by wrinkling on the compression side of the beam, and these at stresses of about 15 tons per square inch, *i.e.*, well above the proportional elastic limit In all the other cases failure occurred on the tension side.

Table IV.

Fairbairn. No. of Experiment	t/R.	Maximum stress at failure.	Type of failure.
2	0·00623	14·92	Compression
6	0·00636	15·1	Tension.
1	0·00673	14·92	Compression.
4	0·00715	14·49	Tension
7	0·00733	10·04	,,
9	0·00747	11·2	,,
8	0·0112	10·1	,,
5	0·0132	13·4	,,
3	0·0214	15·8	,,

In this connection the tubes of the Chepstow Bridge, built by Brunel* about 1850, are also very instructive. The compression member of the bridge is a tube 9 feet diameter and ⅝ inch thick, *i.e.*, $t/R = 0{\cdot}0116$. Diaphragms are provided every 26 feet to preserve the shape. The distance, however, is so great compared with the wave-length of a lobe that they have no strengthening effect The whole truss with transverse girder and roadway weighs 460 tons, and this was assembled on the river bank and tested with 770 tons before being built in position. Under this loading the stress must have been of the order of 4¾ tons per square inch. Brunel had, of course, the experiments of Fairbairn on which to base his design. It is significant that Lilly's formula would give 2·45 tons per square inch as the collapsing load of a tube of this value of t/R, and that Messrs. Barling and Webb's formula would give 2·6 tons as the value of the stress at which wrinkling occurs. Further comment as to the insufficiency of these formulæ is unnecessary.

---

* I. Brunel, 'Life of Brunel' (1870).

The author desires to express his thanks to Sir John Petavel, F.R.S., for his constant help and encouragement when some of the work was being carried out in his laboratory when he was Professor of Engineering at Manchester; to the Superintendent of the Royal Aircraft Establishment and the Air Ministry for permission to publish some of the results obtained whilst the author was at the Royal Aircraft Establishment; to his assistants at the Royal Aircraft Establishment, Mr. I. J. Gerard, M.Sc., and Miss Pettifer, B Sc.; to his present research assistant, Mr. A. J. Newport, B.Sc., for assistance in carrying out some of the experiments, and especially to his friend, Mr R V. Southwell, F.R S., for many helpful discussions in problems connected with this and other investigations.

On the Stability of Running of Locomotives.

By F. W. Carter, M.A., Sc.D., M Inst C.E., M.I.E.E.

(Communicated by A. E. H. Love, F.R.S.—Received July 25, 1928.)

1. *Introduction.*

In a paper entitled "The Electric Locomotive," read before the Institution of Civil Engineers in 1916,* the author gave the beginnings of a rational discussion of the riding qualities of electric locomotives, as affected by their fundamental type. It was shown that, aside from all questions of unevenness of track or of the nature of the springing and equalizing arrangements, there might exist in a locomotive an inherent tendency to deviate from a motion of pure progression, which, although perhaps within the control of the wheel flanges and therefore deemed unobjectionable on good track, might give reason for anxiety if poor track were encountered or other external circumstances conspired with inherent tendencies to bring about a condition involving danger of derailment. It is the intention, in the present paper, to discuss the matter further with the view of determining to what extent inherent riding qualities of a design can be brought within the realm of rational investigation.

* 'Proc. Inst. C.E.,' vol. 201, p. 221.

In applying mathematical analysis to such a subject as the present, it is necessary to assume that all parts fulfil their intentions; though, actually, few do so perfectly. The track is assumed smooth, uniform and level; the clearances required by mechanical considerations, except in so far as they enter into the problem, are assumed to be zero; lubricated slides are assumed frictionless; and so on. It must be postulated that a tendency shown by a machine in which the parts fulfil their intentions will, in general, evince itself in some form in such machines as can actually be constructed. The observed phenomena are likely to reflect the conditions in being less definite, less regular, and more catastrophic than the analysis would indicate; and the results of the present analysis are not expected to possess metrical value of a high order. This, however, is in the nature of the subject and does not detract from the value of the study of tendencies, and the importance to the designer of a knowledge of conditions tending to produce stable running.

2. *General.*

From the present point of view, a locomotive consists of one or more rigid frames, each symmetrical about a longitudinal central plane, supported by springs from a group of axles, and carrying appropriate driving mechanism. The purpose of a frame is to constrain the axles, so that they keep in certain transverse vertical planes fixed in the frame, with the centre point of each on the longitudinal plane of symmetry, whilst permitting to the axles freedom of rotation and a certain freedom of motion in their several vertical planes The locomotive may consist of one or more units, and each unit may comprise a single main frame, or a main frame with one or two auxiliary frames,* these being capable of prescribed movement with respect to the main frame, and biassed towards symmetrical positions by prescribed forces. Each axle bears a pair of wheels, rigidly attached to it; and the treads of the wheels are coned slightly, tapering outwards, (1 in 20), the rails on which they run being generally canted inwards to a like extent.

In the performance of its functions the frame imposes constraint on the axles and vice versa, giving rise to those forces at the wheel-treads which are the ultimate guiding forces of the locomotive. The consideration of these forces is the characteristic and fundamental feature of the present study.

---

* The arrangement of a locomotive unit is, in this country, usually symbolised by a group of three figures, representing successively the number of wheels in the fore auxiliary truck. the main truck, and the rear auxiliary truck.

3. *On the Action between Wheel and Rail.*

The area of contact between wheel and rail has a shape depending on the initial form of the surfaces in contact. When the rail is new, the radius of the transverse section of its head or running surface, is usually smaller than the radius of a wheel; and the contact area is bounded by an ellipse having its major axis in the direction of the rail. The head is flattened by use however, and with a worn rail the major axis of the contact ellipse is transverse to the rail. When the radius of the section of the rail head is equal to that of the wheel, the contact area is circular, having a diameter of the order of 14 mm.[*]

As the wheel rolls, the area of contact travels along the rail and about the periphery of the wheel. If the force between wheel and rail is normal to the contact area, points which come into contact remain so engaged throughout the period of contact. If, however, the force has a component tangential to the contact area, the engagement or meshing of the surfaces is confined to an initial portion of the period of contact, the final portion being characterized by a relative slipping of the surfaces against the tangential force[†] In this case, moreover, the engagement is not between unstrained, or equally strained, surfaces, but between surfaces which are oppositely strained, the tangential force causing extension of the surface of one member, opposed by contraction of the other, and vice versa. In consequence of this, the rolling displacement of the wheel is, under the influence of the tangential force, accompanied by a creeping displacement in the general direction of the force. The latter displacement is proportional to the former, and is otherwise a function of the tangential force.

Defining "creepage" as the ratio between the creeping and rolling displacements, and designating it by γ, the tangential force acting at the tread of the wheel may be written[‡]

$$F = -f\gamma. \tag{1}$$

The value of f has been determined by the author (*loc. cit*) for a certain case in which the tangential force is in the direction of rolling, its value for steel wheels and rails being about

$$f = 93\,[alw]^{\frac{1}{2}}\,[1 + (1 - q)^{\frac{1}{2}}], \tag{2}$$

---

[*] See Love's 'Elasticity,' par. 138. The figure given is based on a pressure of 8,000 kg., and a radius of 600 mm.

[†] See "On the Action of a Locomotive Driving Wheel," by F. W. Carter, 'Roy. Soc. Proc.,' A, vol. 112, p. 151 (1926).

[‡] The notation used is given in an Appendix, pp. 609–11.

in which a is the radius of the wheel in mm.; l is a linear magnitude, representing the effective length of the contact area transverse to the rail in mm.; w is the total pressure between the surfaces, *i.e.*, the weight borne by the wheel, in kilograms; q is the ratio of the tangential force to the maximum that could be sustained without skidding or bodily slipping; and f is in kilograms.* It will be seen that f increases in the ratio 1 : 2 as F falls from its maximum value to zero.

The effective value of f is a matter of importance in the present investigation. The formula given above is insufficient on two counts, viz., the effective value of l is unknown, and probably varies with the condition of the track, and the effective value of q is matter for conjecture. These defects are properly taken into account by reference to general experience in the subject, or to experiment devised to determine a quantity, A, such that

$$f = \mathrm{A}\,(aw)^{\frac{1}{2}}. \tag{3}$$

In the absence of such empirical knowledge, we herein assume

$$\mathrm{A} = 800,$$

which may be interpreted as corresponding to $l = 20$ mm. and $q = 0{\cdot}15$.

Whether the creepage per unit tangential force is the same when the force acts at right angles to the direction of rolling as when the directions coincide or oppose, is matter for conjecture. The most that can be asserted with confidence at present is that, depending as it does on the difference in initial strain of the surfaces which engage one another, it is of the same order whatever the direction of the force, and no reason is apparent why its value should not be independent of this direction. Until fuller knowledge is available, accordingly, we assume the creepage coefficient f, given by equations 2 or 3, to be applicable whatever the direction of the tangential force.†

We may remark that a lack of precise knowledge in these matters does not vitiate general conclusions. We have given expression to our ignorance by introducing the quantity A. If we may assume that, for the several wheel-groups of a locomotive, this quantity is sensibly a constant within the rough limits that the nature of the subject renders necessary, the absence of knowledge of its actual effective value only affects absolute results. The curves of figs. 6 to 9 and 11 to 14, for instance, are affected as to scale, but not as to shape.

---

* With lengths in inches and forces in pounds, the coefficient in the equation becomes 3500 instead of 93.

† In the author's book 'Railway Electric Traction' (Arnold), different creepage coefficients were assumed for forces along and transverse to the rails. No new phenomena were, however, indicated.

The extent to which creepage motions are involved in locomotive running will now be considered. The maximum tractive effort of a wheel before skidding is usually of the order of 35 per cent. of the pressure on the tread.* Assume $a = 625$ mm., $l = 25$ mm., $w = 9000$ kg. and $q = 1$, equation 2 gives $f = 1 \cdot 1 \times 10^6$ kg. Hence the creepage is

$$0 \cdot 35 \times 9000/(1 \cdot 1 \times 10^6) = 2 \cdot 86 \times 10^{-3}.$$

If the treads of a pair of wheels are 1·5 m. apart (which corresponds approximately with standard gauge) the combination can be caused to traverse, without skidding, any curve of radius not less than R metres, given by

$$(\mathrm{R} + 0 \cdot 75/(\mathrm{R} - 0 \cdot 75) = (1 + 2 \cdot 86 \,.\, 10^{-3})/(1 - 2 \cdot 86 \,.\, 10^{-3})$$

or

$$\mathrm{R} = 262 \text{ metres}.$$

This is a small radius for a railway curve, and locomotives therefore take normal service curves without skidding the wheels. Again, the pair of wheels could be run on straight track, without skidding, if their respective radii were $625\,(1 + 2 \cdot 86 \times 10^{-3})$ mm. and $625\ (1 - 2 \cdot 86 \times 10^{-3})$ mm. Since the conicity of the wheels is 1 in 20, this difference corresponds with a movement of the pair from the central position of $20 \times 625 \times 2 \cdot 86 \times 10^{-3}$ mm., *i.e.*, 35·8 mm., or a total play of 71·6 mm., which is certainly more than double the movement which the flanges would permit.† It may be concluded that the motions of deviation from pure progression which are possible to a locomotive are of the nature of creeping rather than slipping.

4. *Simple Rolling of Wheels and Axle.*

Fig. 1 represents a pair of railway wheels mounted on axle, running on rails canted λ to correspond with the coning of the wheels. Taking axes of co-ordinates in the plane of the track, with the centre line of the track as x-axis, let (xy) be the co-ordinates of the centre of the axle, and ψ its inclination with the y-axis, in the plan view. The effective radius of the wheel A is

---

* In electric locomotives having independently driven axles, experience has shown that, under ordinary circumstances, a tractive effort of 25 to 30 per cent. of the weight on drivers can be sustained; but the pull of the draw-bar re-distributes the pressure on the several axles, decreasing that on the forward axles, which accordingly limit the tractive effort. The limiting ratio for these is usually at least 35 per cent.

† The play permitted between flanges and rails, as fixed by the Berne Conference, May 15, 1886, is, minimum 15 mm., maximum 35 mm. See 'Traité de Stabilité du Matériel des Chemins de Fer,' par Georges Marié, page 306.

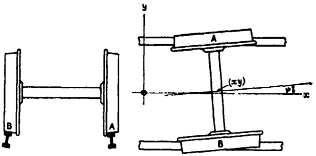


FIG. 1.—Diagram of Wheels and Axle.

$a + \lambda y$, and of the wheel B, $a - \lambda y$. The co-ordinates of the two points of contact are—

$$\text{of A } \{x - b\psi,\ y + b\},\ \text{of B } \{x + b\psi,\ y - b\}.$$

The component displacements are, therefore,

$$\text{of A } \{dx - bd\psi,\ dy\},\ \text{of B } \{dx + bd\psi,\ dy\}, \tag{4}$$

but the components of rolling displacement are,

$$\text{of A } \{(a + \lambda y)d\theta,\ a\psi d\theta\},\ \text{and of B } \{(a - \lambda y)d\theta,\ a\psi d\theta\}; \tag{5}$$

and, accordingly, for simple rolling,

$$dx = ad\theta, \quad dy = a\psi d\theta, \quad bd\psi = -\lambda y d\theta.$$

Thus

$$ab\frac{d^2\psi}{dx^2} + \lambda\psi = 0. \tag{6}$$

The motion is, therefore, sinuous and repeats in distance,

$$X = 2\pi\sqrt{ab/\lambda}. \tag{7}$$

5. *Constrained Rolling of Wheels and Axle.*

It has been shown that if appropriate constraint be applied to the axle, the rolling motion is accompanied by creeping. The components of creepage displacement are, from 4 and 5 above,

$$\left.\begin{array}{l} \text{of A } \{dx - b\,d\psi - (a + \lambda y)\,d\theta,\quad dy - a\psi d\theta\} \\ \text{of B } \{dx + b\,d\psi - (a - \lambda y)d\theta,\quad dy - a\psi d\theta\} \end{array}\right\}. \tag{8}$$

The components of creepage are

$$\left.\begin{array}{l} \text{of A } \left\{\dfrac{dx}{a\,d\theta} - 1 - \left(\dfrac{b}{a}\dfrac{d\psi}{d\theta} + \lambda\dfrac{y}{a}\right),\ \dfrac{dy}{a\,d\theta} - \psi\right\} \\ \text{of B } \left\{\dfrac{dx}{a\,d\theta} - 1 + \left(\dfrac{b}{a}\dfrac{d\psi}{d\theta} + \lambda\dfrac{y}{a}\right),\ \dfrac{dy}{a\,d\theta} - \psi\right\} \end{array}\right\}. \tag{9}$$

The component forces at the wheel-treads are, accordingly,

$$\left.\begin{array}{l} \text{at A} \left\{-f\left[\dfrac{dx}{ad\theta} - 1 - \left(\dfrac{b}{a}\,\dfrac{d\psi}{d\theta} + \lambda\dfrac{y}{a}\right)\right],\; -f\left[\dfrac{dy}{ad\theta} - \psi\right]\right\} \\ \text{at B} \left\{-f\left[\dfrac{dx}{ad\theta} - 1 + \left(\dfrac{b}{a}\,\dfrac{d\psi}{d\theta} + \lambda\dfrac{y}{a}\right)\right],\; -f\left[\dfrac{dy}{ad\theta} - \psi\right]\right\} \end{array}\right\}. \tag{10}$$

This system of forces is equivalent to—

1. A longitudinal force at the centre of the axle, of value,

$$X_1 = -2f\left(\frac{dx}{ad\theta} - 1\right) \tag{11}$$

2. A transverse force at the centre of the axle, of value,

$$Y_1 = -2f\left(\frac{dy}{ad\theta} - \psi\right). \tag{12}$$

3 A couple about a vertical axis, having moment

$$G_1 = -2f\left(\frac{b^2}{a}\,\frac{d\psi}{d\theta} + \lambda\,\frac{b}{a}\,y\right). \tag{13}$$

If the system of forces be referred to some fixed point, Q (fig. 2), on the centre line of the locomotive or truck (in the plan view), having co-ordinates

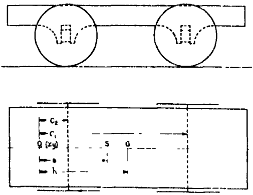


Fig 2.—Diagram of Truck.

(xy), and at distance c behind the axle, the components of the system become, to the degree of approximation required,

$$X = -2f\left[\frac{dx}{ad\theta} - 1\right], \tag{14}$$

$$Y = -2f\left[\frac{dy}{ad\theta} + c\,\frac{d\psi}{ad\theta} - \psi\right], \tag{15}$$

$$G = -2f\left[(b^2 + c^2)\,\frac{d\psi}{ad\theta} - \left(1 - \lambda\frac{b}{a}\right)c\psi + c\,\frac{dy}{ad\theta} + \lambda\frac{b}{a}\,y\right]. \tag{16}$$

Since f is large compared with the greatest possible value of X, $dx/ad\theta$ differs but little from unity, and dx may be substituted for $ad\theta$ in equations 15 and 16.

6. *Equations of Motion.*

Let the centre of gravity of the portion of the locomotive affected by the group of wheels under consideration be distant, h, ahead of the reference point Q (fig. 2). Let m be the mass of the part, and mk^2 its moment of inertia about the vertical through the centre of gravity. Let X′Y′G′ be components and moment of other forces acting on the part, such as those produced by auxiliary trucks, these being reduced to some point S on the axis of symmetry of the locomotive, distant s ahead of the reference point for the wheel-forces and motion. The equations of motion of the part are—

$$m\frac{d^2x}{dt^2} = \mathrm{X}' + \Sigma\mathrm{X}, \tag{17}$$

$$m\left(\frac{d^2y}{dt^2} + h\frac{d^2\psi}{dt^2}\right) = \mathrm{Y}' + \Sigma\mathrm{Y}, \tag{18}$$

$$mk^2\frac{d^2\psi}{dt^2} = \mathrm{G}' + \Sigma\mathrm{G} - h\left(\mathrm{Y}' + \Sigma\mathrm{Y}\right) + s\mathrm{Y}'. \tag{19}$$

Equation 19 may be replaced by—

$$m\left(h^2 + k^2\right)\frac{d^2\psi}{dt^2} + mh\frac{d^2y}{dt^2} = \mathrm{G}' + \Sigma\mathrm{G} + s\mathrm{Y}'. \tag{20}$$

In these equations, the summations extend over the several axles of the group under consideration Equation 17 has no particular interest in regard to matters now under consideration, being concerned with the normal progression of the locomotive. We do not take account of the effect of the forward acceleration of the locomotive on the phenomenon of deviation. Accordingly, if V is the velocity of progression, it is permissible to write—

$$\frac{d^2}{dt^2} = \mathrm{V}^2\frac{d^2}{dx^2}$$

and, taking account of the remark which follows equation 16, equations 18 and 20 reduce to

$$m\mathrm{V}^2\left[\frac{d^2y}{dx^2} + h\frac{d^2\psi}{dx^2}\right] + 2\Sigma f\left[\frac{dy}{dx} + c\frac{d\psi}{dx} - \psi\right] = \mathrm{Y}', \tag{21}$$

$$m\mathrm{V}^2\left[h\frac{d^2y}{dx^2} + \left(h^2 + k^2\right)\frac{d^2\psi}{dx^2}\right] + 2\Sigma f\left[c\frac{dy}{dx} + \lambda\frac{b}{a}y + \left(b^2 + c^2\right)\frac{d\psi}{dx} - \left(1 - \lambda\frac{b}{a}\right)c\psi\right] = \mathrm{G}' + s\mathrm{Y}'. \tag{22}$$

These are the equations which determine the deviation of the motion from normal progression. Since, in applying them to problems of particular locomotives, expressions of considerable length and complexity arise, it is expedient to adopt a notation which reduces the amount of writing to the minimum. The notation used is, for convenience, given in an appendix to the paper. In terms of this notation, writing D for d/dx, equations 21 and 22 become—

$$[\mathbf{a}D^2 + D]y + [\mathbf{a}hD^2 + \mathbf{c}D - 1]\,\psi = \tfrac{1}{2}\mathbf{k}Y', \tag{23}$$

$$[\mathbf{a}hD^2 + \mathbf{c}D + \mathbf{s}]\,y + [\mathbf{a}\,(h^2 + k^2)\,D^2 + \mathbf{d}^2D + \mathbf{b} - \mathbf{c}]\,\psi = \tfrac{1}{2}\mathbf{k}\,(G' + sY'). \tag{24}$$

7. *Locomotives of One Axle-Group*

For the locomotive running alone, $Y' = 0$, $G' = 0$: equations 23 and 24 therefore give—

$$\{\mathbf{a}^2k^2D^4 + \mathbf{a}\,[h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c}]\,D^3 + [\mathbf{d}^2 - \mathbf{c}^2 + \mathbf{a}\,(\mathbf{b} - \mathbf{c} + h - \mathbf{s}h)]\,D^2 + [\mathbf{b} - \mathbf{sc}]\,D + \mathbf{s}\}\,\psi = 0. \tag{25}$$

This equation indicates stable motion of progression if*

$$\mathbf{b} - \mathbf{sc} > \frac{\mathbf{a}}{\mathbf{d}^2 - \mathbf{c}^2}\left\{(h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c})\mathbf{s} - (\mathbf{b} - \mathbf{c} + h - \mathbf{s}h)\,(\mathbf{b} - \mathbf{sc}) + \frac{k^2\,(\mathbf{b} - \mathbf{sc})^2}{h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c}}\right\}. \tag{26}$$

The condition $\mathbf{b} - \mathbf{sc} > 0$ implies, as shown in the Appendix, that the wheels of the several axles differ in diameter, and that the smaller wheels are ahead. This is not usual in a locomotive having a fixed wheel-base, but only on these conditions is stable running possible, and then only to the speed limit given by the inequality 26.

A numerical example may be of interest as showing the order of quantities involved. Take dimensions and weights as shown in fig. 3, standard gauge ($b = 0{\cdot}75$ m.), the truck running with small wheels ahead.

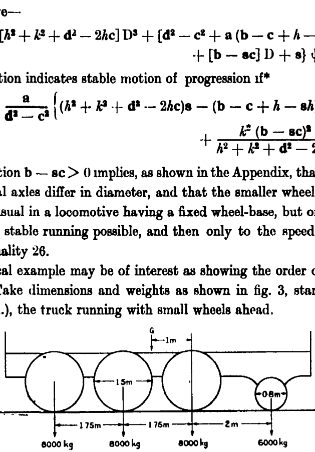


FIG. 3.—Truck carried on rigid wheel-base.

---

* For conditions of stability of a motion represented by a linear differential equation having constant coefficients, see Routh's 'Advanced Rigid Dynamics,' pars. 290-307.

For the front axle $f = 1\cdot24 \times 10^6$, for the three rear axles $f = 1\cdot96 \times 10^6$, hence $\mathbf{f} = 7\cdot12 \times 10^6$. We take the reference point at G, accordingly,

$$\mathbf{c} = -0\cdot0969 \qquad \mathbf{s} = 0\cdot0576 \qquad \mathbf{b} = 0\cdot0180$$

$$\mathbf{b} - \mathbf{c} = 0\cdot1149 \qquad \mathbf{b} - \mathbf{sc} = 0\cdot02354$$

$$\mathbf{d}^2 = 4\cdot281 \qquad \mathbf{d}^2 - \mathbf{c}^2 = 4\cdot271.$$

$$k^2 = 1 \qquad h = 0 \qquad m = 60{,}000 \text{ kg}.$$

$$\mathbf{a} = 4\cdot296 \times 10^{-4}\mathrm{V}^2.$$

The inequality 26 therefore gives

$$0\cdot02354 > \frac{4\cdot296 \times 10^{-4}\,\mathrm{V}^2}{4\cdot271}\{0\cdot477 - 0\cdot0027 + 0\cdot0003\}.$$

We deduce that the running becomes unstable at a speed of 22·2 metres per second or about 50 miles per hour. The three terms within the bracket have been kept separate in order to show that the first alone is of importance: the second arises from the coefficient of $\mathbf{a}$ in the D^2 term in equation 25, and the third from the first term in this equation. We might, accordingly, without sensible error have written for equation 25

$$\{\mathbf{a}\,(h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c})\,\mathrm{D}^3 + (\mathbf{d}^2 - \mathbf{c}^2)\,\mathrm{D}^2 + (\mathbf{b} - \mathbf{sc})\,\mathrm{D} + \mathbf{s}\}\,\psi = 0. \quad (27)$$

Moreover, within the limits imposed by practicability, we can readily obtain an approximate algebraic solution of this equation. Assume

$$\psi = \psi_0\, e^{\zeta x} = \psi_0\, e^{(\xi + i\eta)x},$$

with $\zeta = i\eta$ as a first approximation, where

$$\eta = \left(\frac{\mathbf{s}}{\mathbf{d}^2 - \mathbf{c}^2}\right)^{\frac{1}{2}}. \quad (28)$$

Proceeding to a next approximation, we obtain

$$\xi = \frac{\mathbf{a}\,(h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c})}{2\,(\mathbf{d}^2 - \mathbf{c}^2)^2} - \frac{(\mathbf{b} - \mathbf{sc})}{2\,(\mathbf{d}^2 - \mathbf{c}^2)}. \quad (29)$$

The condition for stability, viz., that ξ should be negative, agrees with 26, with the omission of the last two terms within the brackets. We note that the motion is periodic in distance

$$\mathrm{X} = 2\pi\left(\frac{\mathbf{d}^2 - \mathbf{c}^2}{\mathbf{s}}\right)^{\frac{1}{2}}. \quad (30)$$

This is, of course, independent of the reference point.*

---

* See Appendix.

Since the example given above does not refer to a normal type of locomotive of one axle group, we discuss briefly the more normal type shown in fig. 4.

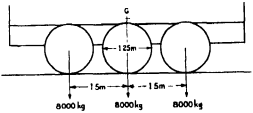


Fig. 4.—Locomotive carried on rigid wheel-base (0–6–0).

Here, taking the centre as reference point for the motion,

$$f = 1{\cdot}79 \times 10^6 \qquad \mathbf{f} = 5{\cdot}37 \times 10^6 \qquad m = 48{,}000 \text{ kg.}$$

$$\mathbf{a} = 4{\cdot}56 \times 10^{-4}\,\mathrm{V}^2 \qquad \mathbf{b} = 0, \quad \mathbf{c} = 0,$$

$$\mathbf{s} = 0{\cdot}06 \qquad \mathbf{d}^2 = 2{\cdot}062 \qquad h = 0, \quad k = 1{\cdot}2,$$

whence

$$\eta = 0{\cdot}1705 \qquad \xi = 1{\cdot}126 \times 10^{-5}\,\mathrm{V}^2$$

$$\mathrm{X} = 36{\cdot}8 \text{ m} \qquad \xi\mathrm{X} = 4{\cdot}15 \times 10^{-4}\,\mathrm{V}^2$$

Thus, at V = 15 m.p.s., the amplitude of deviation increases by about 10 per cent in each complete period; at V = 20 m.p.s., the increase is about 18 per cent, whilst at V = 25 m.p.s., it is about 30 per cent.

The locomotive carried on a single axle-group is much used in working freight trains, but is not employed for high-speed running on account of proclivities indicated in the foregoing discussion. The hauling of a train or tender has some tendency to stabilize the running,* but the effect is not great, being of the second order. The use of side buffers (as distinguished from centre buffers) produces a tendency of the first order making for stable running, as long as they remain in compression.

The periodic digressive motion, at least in its more violent manifestation, is known as "nosing." It is not, properly speaking, an oscillation, but is the result of instability of a pure motion of progression. In the ordinary locomotive of the type, the instability is to be attributed entirely to the inertia of the body, but locomotives having unequal wheels, in pairs, may, as has been shown, have stability or instability independently of the inertia, according to the location of the wheels. Nosing has in its origin nothing to do with the spring or equalizing systems, although it is likely to become noticeable, if not to be actually enhanced, at such speed that the period of the motion, in time, agrees with a natural period of oscillation of the locomotive. The periodic distance is, as has been shown, considerably greater than the length of a rail, and nosing is therefore unaffected by the periodicity in the track.

---

* 'Proc. Inst. C.E.,' vol. 201, p. 251 (1916).

8. *Locomotives of Two Axle-Groups.*

The locomotives now to be discussed have, in general, one group of axles associated with the chief mass of the locomotive, and this is designated the main group; the other group constitutes an auxiliary truck, and the associated parts are usually of comparatively small mass. The auxiliary truck pivots about a point on the centre line of the main frame, the pivot being sometimes situated at a definite point in the truck itself, as in the Bissel, and sometimes permitted to traverse the truck, as in the bogie The design of locomotives in respect of these matters has hitherto been determined principally with reference to the need for steering the locomotive on curves, and providing a flexible wheel-base over which to distribute its weight; but since it has much influence on the running qualities now under consideration, the problems should be correlated

The auxiliary truck is, in general, biassed towards a symmetrical central position by means of springs, or equivalent devices, which apply the steering forces to the main frame. In many locomotives the centring devices are so designed that considerable force is necessary to displace the truck from the position of symmetry, although, once displaced, the variation of the force is comparatively small, or even zero. It is not practicable to take account of this discontinuity, and the centring forces will, in all cases, be assumed proportional to the displacement from the position of symmetry.

We take the pivotal point on the main frame as reference point, both for the motion and for the forces between the two elements. We use dashed letters to distinguish auxiliary from main trucks. We assume a centring force, $-2\mathrm{Y}(y-y')$, to act on the main truck at the pivot, and an aligning couple, $-2\mathrm{G}(\psi-\psi')$, to act about the pivot, thus including all cases in one formula. The equations of devious motion are (see equations 23 and 24),

$$(\mathbf{a}\mathrm{D}^2+\mathrm{D})y+(\mathbf{a}\boldsymbol{h}\mathrm{D}^2+\mathbf{c}\mathrm{D}-1)\psi+\mathbf{k}\mathrm{Y}(y-y')=0, \tag{31}$$

$$(\mathbf{a}'\mathrm{D}^2+\mathrm{D})y'+(\mathbf{a}'\boldsymbol{h}'\mathrm{D}^2+\mathbf{c}'\mathrm{D}-1)\psi'+\mathbf{k}'\mathrm{Y}(y'-y)=0, \tag{32}$$

$$(\mathbf{a}\boldsymbol{h}\mathrm{D}^2+\mathbf{c}\mathrm{D}+\mathbf{s})y+[\mathbf{a}(\boldsymbol{h}^2+\boldsymbol{k}^2)\mathrm{D}^2+\mathbf{d}^2\mathrm{D}+\mathbf{b}-\mathbf{c}]\psi+\mathbf{k}\mathrm{G}(\psi-\psi')=0, \tag{33}$$

$$(\mathbf{a}'\boldsymbol{h}'\mathrm{D}^2+\mathbf{c}'\mathrm{D}+\mathbf{s}')y'+[\mathbf{a}'(\boldsymbol{h}'^2+\boldsymbol{k}'^2)\mathrm{D}^2+\mathbf{d}'^2\mathrm{D}+\mathbf{b}'-\mathbf{c}']\psi' \\ +\mathbf{k}'\mathrm{G}(\psi'-\psi)=0 \tag{34}$$

These give, as the equation which determines stability,

$$\begin{vmatrix} \mathbf{a}D^2+D+\mathbf{k}Y & -\mathbf{k}Y & \mathbf{a}hD^2+\mathbf{c}D-1 & 0 \\ -\mathbf{k}'Y & \mathbf{a}'D^2+D+\mathbf{k}'Y & 0 & \mathbf{a}'h'D^2+\mathbf{c}'D-1 \\ \mathbf{a}hD^2+\mathbf{c}D+\mathbf{s} & 0 & \begin{array}{r}\mathbf{a}(h^2+k^2)D^2+\mathbf{d}^2D \\ +\mathbf{b}-\mathbf{c}+\mathbf{k}G\end{array} & -\mathbf{k}G \\ 0 & \mathbf{a}'h'D^2+\mathbf{c}'D+\mathbf{s}' & -\mathbf{k}'G & \begin{array}{r}\mathbf{a}'(h'^2+k'^2)D^2+\mathbf{d}'^2D \\ +\mathbf{b}'-\mathbf{c}'+\mathbf{k}'G\end{array} \end{vmatrix} = 0. \qquad (35)$$

This is the characteristic equation of the running of locomotives of two axle-groups, in its general form. The expansion of the determinant is long; but, since it is symmetrical as regards the trucks, it is not difficult.

It is expedient to divide the terms into four categories, viz, those independent of Y and G, those involving Y only, those involving G only, and those involving both Y and G. The terms of the first category form the product of two expressions similar to equation 25. They are as follows :—

$$\begin{aligned} &\mathbf{a}^2\mathbf{a}'^2k^2k'^2D^8 + \mathbf{aa}'\{\mathbf{a}k^2(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}') + \mathbf{a}'k'^2(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})\}D^7 \\ &+\{\mathbf{a}^2k^2(\mathbf{d}'^2-\mathbf{c}'^2)+\mathbf{a}'^2k'^2(\mathbf{d}^2-\mathbf{c}^2)+\mathbf{aa}'(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}') \\ &\qquad +\mathbf{aa}'[\mathbf{a}k^2(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')+\mathbf{a}'^2k'^2(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)]\}D^6 \\ &+\{\mathbf{a}(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})(\mathbf{d}'^2-\mathbf{c}'^2)+\mathbf{a}'(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')(\mathbf{d}^2-\mathbf{c}^2) \\ &\qquad +\mathbf{aa}'[(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')+(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)] \\ &\qquad +\mathbf{a}^2k^2(\mathbf{b}'-\mathbf{s}'\mathbf{c}')+\mathbf{a}'^2k'^2(\mathbf{b}-\mathbf{sc})\}D^5 \\ &+\{(\mathbf{d}^2-\mathbf{c}^2)(\mathbf{d}'^2-\mathbf{c}'^2)+\mathbf{a}(\mathbf{d}'^2-\mathbf{c}'^2)(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)+\mathbf{a}'(\mathbf{d}^2-\mathbf{c}^2)(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h') \\ &\qquad +\mathbf{a}(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})(\mathbf{b}'-\mathbf{s}'\mathbf{c}')+\mathbf{a}'(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')(\mathbf{b}-\mathbf{sc}) \\ &\qquad +\mathbf{a}^2k^2\mathbf{s}'+\mathbf{a}'^2k'^2\mathbf{s}+\mathbf{aa}'(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h'\}D^4 \\ &+\{(\mathbf{d}^2-\mathbf{c}^2)(\mathbf{b}'-\mathbf{s}'\mathbf{c}')+(\mathbf{d}'^2-\mathbf{c}'^2)(\mathbf{b}-\mathbf{sc})+\mathbf{a}(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})\mathbf{s}' \\ &\qquad +\mathbf{a}'(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')\mathbf{s}+\mathbf{a}(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)(\mathbf{b}'-\mathbf{s}'\mathbf{c}') \\ &\qquad +\mathbf{a}'(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')(\mathbf{b}-\mathbf{sc})\}D^3 \\ &+\{(\mathbf{d}^2-\mathbf{c}^2)\mathbf{s}'+(\mathbf{d}'^2-\mathbf{c}'^2)\mathbf{s}+(\mathbf{b}-\mathbf{sc})(\mathbf{b}'-\mathbf{s}'\mathbf{c}')+\mathbf{a}(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)\mathbf{s}' \\ &\qquad +\mathbf{a}'(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')\mathbf{s}\}D^2 \\ &+\{(\mathbf{b}-\mathbf{sc})\mathbf{s}'+(\mathbf{b}'-\mathbf{s}'\mathbf{c}')\mathbf{s}\}D+\mathbf{ss}'. \qquad (36) \end{aligned}$$

The quantities Y and G are each involved to the first degree only. The terms involving Y only, omitting the factor Y, are as follows :—

$$\begin{aligned}
&\mathbf{kaa}'^2k'^2(h^2+k^2)+\mathbf{k}'\mathbf{a}'\mathbf{a}^2k^2(h'^2+k'^2)\}\,\mathrm{D}^6\\
&\quad+\{\mathbf{k}[\mathbf{a}'^2k'^2\mathbf{d}^2+\mathbf{aa}'(h^2+k^2)(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')]\\
&\qquad+\mathbf{k}'[\mathbf{a}^2k^2\mathbf{d}'^2+\mathbf{aa}'(h'^2+k'^2)(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})]\}\,\mathrm{D}^5\\
&\quad+\{\mathbf{k}[\mathbf{a}(h^2+k^2)(\mathbf{d}'^2-\mathbf{c}'^2)+\mathbf{a}'\mathbf{d}^2(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')\\
&\qquad+\mathbf{aa}'(h^2+k^2)(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')+\mathbf{a}'^2k'^2(\mathbf{b}-\mathbf{c})]\\
&\qquad+\mathbf{k}'[\mathbf{a}'(h'^2+k'^2)(\mathbf{d}^2-\mathbf{c}^2)+\mathbf{ad}'^2(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})\\
&\qquad+\mathbf{aa}'(h'^2+k'^2)(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)+\mathbf{a}^2k^2(\mathbf{b}'-\mathbf{c}')]\}\,\mathrm{D}^4\\
&\quad+\{\mathbf{kd}^2(\mathbf{d}'^2-\mathbf{c}'^2)+\mathbf{k}'\mathbf{d}'^2(\mathbf{d}^2-\mathbf{c}^2)\\
&\qquad+\mathbf{k}[\mathbf{a}(h^2+k^2)(\mathbf{b}'-\mathbf{s}'\mathbf{c}')+\mathbf{a}'\mathbf{d}^2(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')+\mathbf{a}'(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')(\mathbf{b}-\mathbf{c})]\\
&\qquad+\mathbf{k}'[\mathbf{a}'(h'^2+k'^2)(\mathbf{b}-\mathbf{sc})+\mathbf{ad}'^2(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)+\mathbf{a}(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})(\mathbf{b}'-\mathbf{c}')]\}\,\mathrm{D}^3\\
&\quad+\{\mathbf{k}[(\mathbf{d}'^2-\mathbf{c}'^2)(\mathbf{b}-\mathbf{c})+\mathbf{d}^2(\mathbf{b}'-\mathbf{s}'\mathbf{c}')]+\mathbf{k}'[(\mathbf{d}^2-\mathbf{c}^2)(\mathbf{b}'-\mathbf{c}')+\mathbf{d}'^2(\mathbf{b}-\mathbf{sc})]\\
&\qquad+\mathbf{k}[\mathbf{a}(h^2+k^2)\mathbf{s}'+\mathbf{a}'(\mathbf{b}-\mathbf{c})(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')]\\
&\qquad+\mathbf{k}'[\mathbf{a}'(h'^2+k'^2)\mathbf{s}+\mathbf{a}(\mathbf{b}'-\mathbf{c}')(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)]\}\,\mathrm{D}^2\\
&\quad+\{\mathbf{k}[\mathbf{d}^2\mathbf{s}'+(\mathbf{b}-\mathbf{c})(\mathbf{b}'-\mathbf{s}'\mathbf{c}')]+\mathbf{k}'[\mathbf{d}'^2\mathbf{s}+(\mathbf{b}'-\mathbf{c}')(\mathbf{b}-\mathbf{sc})]\}\,\mathrm{D}\\
&\qquad+\mathbf{k}(\mathbf{b}-\mathbf{c})\mathbf{s}'+\mathbf{k}'(\mathbf{b}'-\mathbf{c}')\mathbf{s}.
\end{aligned}\tag{37}$$

The terms involving G only, omitting the factor G, are as follows :—

$$\begin{aligned}
&(\mathbf{ka}'k'^2+\mathbf{k}'\mathbf{a}k^2)\,\mathbf{aa}'\,\mathrm{D}^6\\
&\quad+\{\mathbf{k}[\mathbf{aa}'(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')+\mathbf{a}'^2k'^2]\\
&\qquad+\mathbf{k}'[\mathbf{aa}'(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})+\mathbf{a}^2k^2]\}\,\mathrm{D}^5\\
&\quad+\{\mathbf{k}[\mathbf{a}(\mathbf{d}'^2-\mathbf{c}'^2)+\mathbf{a}'(h'^2+k'^2+\mathbf{d}'^2-2h'\mathbf{c}')+\mathbf{aa}'(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')]\\
&\qquad+\mathbf{k}'[\mathbf{a}'(\mathbf{d}^2-\mathbf{c}^2)+\mathbf{a}(h^2+k^2+\mathbf{d}^2-2h\mathbf{c})+\mathbf{aa}'(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)]\}\,\mathrm{D}^4\\
&\quad+\{\mathbf{k}[\mathbf{d}'^2-\mathbf{c}'^2+\mathbf{a}'(\mathbf{b}'-\mathbf{c}'+h'-\mathbf{s}'h')+\mathbf{a}(\mathbf{b}'-\mathbf{s}'\mathbf{c}')]\\
&\qquad+\mathbf{k}'[\mathbf{d}^2-\mathbf{c}^2+\mathbf{a}(\mathbf{b}-\mathbf{c}+h-\mathbf{s}h)+\mathbf{a}'(\mathbf{b}-\mathbf{sc})]\}\,\mathrm{D}^3\\
&\quad+\{\mathbf{k}[\mathbf{b}'-\mathbf{s}'\mathbf{c}'+\mathbf{as}']+\mathbf{k}'[\mathbf{b}-\mathbf{sc}+\mathbf{a}'\mathbf{s}]\}\,\mathrm{D}^2\\
&\quad+\{\mathbf{ks}'+\mathbf{k}'\mathbf{s}\}\,\mathrm{D}.
\end{aligned}\tag{38}$$

The terms involving YG, omitting this factor, are :—

$$\begin{aligned}
&\{(\mathbf{ka}' + \mathbf{k}'\mathbf{a})\,(\mathbf{ka}'k'^2 + \mathbf{k}'\mathbf{a}k^2) + \mathbf{kk}'\,\mathbf{aa}'\,(h - h')^2\}\,\mathrm{D}^4 \\
&\quad + \{\mathbf{ka}'\,[\mathbf{k}\,(h'^2 + k'^2 + \mathbf{d}'^2 - 2h'\mathbf{c}') + \mathbf{k}'\,(h'^2 + k'^2 + \mathbf{d}^2 - 2h'\mathbf{c})] \\
&\qquad + \mathbf{k}'\mathbf{a}\,[\mathbf{k}'\,(h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c}) + \mathbf{k}\,(h^2 + k^2 + \mathbf{d}'^2 - 2h\mathbf{c}')]\}\,\mathrm{D}^3 \\
&\quad + \{(\mathbf{k} + \mathbf{k}')\,(\mathbf{kd}'^2 + \mathbf{k}'\mathbf{d}^2) - (\mathbf{kc}' + \mathbf{k}'\mathbf{c})^2 \\
&\qquad + \mathbf{ka}'\,[\mathbf{k}\,(\mathbf{b}' - \mathbf{c}' + h' - \mathbf{s}'h') + \mathbf{k}'\,(\mathbf{b} - \mathbf{c} + h' - \mathbf{s}h')] \\
&\qquad + \mathbf{k}'\mathbf{a}\,[\mathbf{k}'\,(\mathbf{b} - \mathbf{c} + h - \mathbf{s}h) + \mathbf{k}\,(\mathbf{b}' - \mathbf{c}' + h - \mathbf{s}'h)]\}\,\mathrm{D}^2 \\
&\quad + \{(\mathbf{k} + \mathbf{k}')\,(\mathbf{kb}' + \mathbf{k}'\mathbf{b}) - (\mathbf{ks}' + \mathbf{k}'\mathbf{s})\,(\mathbf{kc}' + \mathbf{k}'\mathbf{c})\}\,\mathrm{D} \\
&\qquad + (\mathbf{k} + \mathbf{k}')\,(\mathbf{ks}' + \mathbf{k}'\mathbf{s}). \qquad (39)
\end{aligned}$$

The expressions are given in full, but in most problems simplification is possible depending on the nature of the case. Usually all wheels of a group are alike, so that $\mathbf{b} - \mathbf{sc} = 0$ for the group. Terms involving $\mathbf{a}'$, which is proportional to the mass of the auxiliary truck, are often small, though not always negligible. The terms of the highest degree seem usually to have little effect on stability. It, however, adds little labour to form the characteristic equation of the locomotive for any particular case, completely, and with all accuracy that the data permits

We consider first a locomotive of very usual type, having an auxiliary bogie, arranged as shown diagrammatically in fig. 5. The pivotal connection, P, is at the centre of the bogie, in the longitudinal view, and is biassed towards the centre of the transverse view by a force proportional to the displacement ; no aligning couple is imposed on the bogie. The characteristic equation for the locomotive is therefore formed from expressions 36 and 37 only. In the example discussed, for which dimensions and other data are given in fig. 5,

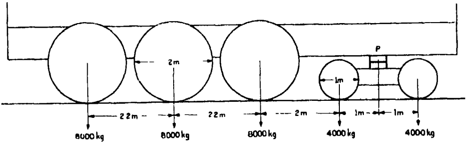


Fig. 5.—Bogie Locomotive (4-6-0 or 0-6-4).

the mass of the main frame and parts associated with it, is taken as 60,000 kg., and that of the auxiliary bogie as 4,000 kg, the wheel pressures being as indicated in the figure. Following are derived and assumed data on which the characteristic equation has been based.

$$\begin{array}{lll}
f = 2{\cdot}263 \times 10^6 & \mathbf{f} = 6{\cdot}788 \times 10^6 & \mathbf{k} = 1{\cdot}473 \times 10^{-7} \\
f' = 1{\cdot}131 \times 10^6 & \mathbf{f}' = 2{\cdot}263 \times 10^6 & \mathbf{k}' = 3\mathbf{k}
\end{array}$$

$$\begin{array}{ll}
\mathbf{c} = -5{\cdot}2 & \mathbf{c}' = 0 \\
\mathbf{s} = 0{\cdot}0375 & \mathbf{s}' = 0{\cdot}075 \\
\mathbf{b} - \mathbf{c} = 5{\cdot}005 & \mathbf{b}' - \mathbf{c}' = 0 \\
\mathbf{d}^2 = 30{\cdot}83 & \mathbf{d}'^2 = 1{\cdot}5625 \\
\mathbf{d}^2 - \mathbf{c}^2 = 3{\cdot}789 & \mathbf{d}'^2 - \mathbf{c}'^2 = 1{\cdot}5625 \\
\mathbf{a} = 4{\cdot}5 \times 10^{-4}\,\mathrm{V}^2 & \mathbf{a}' = 0{\cdot}2\,\mathbf{a} \\
h = -4{\cdot}16 & h' = 0 \\
k^2 = 12 & k'^2 = 0{\cdot}6 \\
h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c} = 16{\cdot}871 & h'^2 + k'^2 + \mathbf{d}'^2 - 2h'\mathbf{c}' = 2{\cdot}162 \\
\mathbf{b} - \mathbf{c} + h - \mathbf{s}h = 1{\cdot}001 & \mathbf{b}' - \mathbf{c}' + h' - \mathbf{s}'h' = 0.
\end{array}$$

The characteristic equation for forward motion, i.e., with bogie ahead, is :—

$$\begin{aligned}
&0{\cdot}288\,\mathbf{a}^4\mathrm{D}^8 + 5{\cdot}595\mathbf{a}^3\mathrm{D}^7 + [26{\cdot}135\mathbf{a}^2 + (0{\cdot}0240 + 5{\cdot}023\mathbf{k}\mathrm{Y})\,\mathbf{a}^3]\,\mathrm{D}^6 \\
&\quad + [28{\cdot}00\mathbf{a} + (0{\cdot}433 + 75{\cdot}74\mathbf{k}\mathrm{Y})\,\mathbf{a}^2]\,\mathrm{D}^5 \\
&\quad + [5{\cdot}921 + (1{\cdot}564 + 139{\cdot}6\mathbf{k}\mathrm{Y})\,\mathbf{a} + (0{\cdot}901 + 0{\cdot}4806\mathbf{k}\mathrm{Y})\,\mathbf{a}^2]\,\mathrm{D}^4 \\
&\quad + [65{\cdot}93\mathbf{k}\mathrm{Y} + (1{\cdot}2815 + 6{\cdot}856\mathbf{k}\mathrm{Y})\,\mathbf{a}]\,\mathrm{D}^3 \\
&\quad + [0{\cdot}3428 + 7{\cdot}820\mathbf{k}\mathrm{Y} + (0{\cdot}0751 + 2{\cdot}211\mathbf{k}\mathrm{Y})\,\mathbf{a}]\,\mathrm{D}^2 \\
&\quad + 2{\cdot}488\mathbf{k}\mathrm{YD} + 0{\cdot}002812 + 0{\cdot}3754\mathbf{k}\mathrm{Y} = 0.
\end{aligned} \tag{40}$$

At low speed ($\mathbf{a} = 0$) this reduces to

$$\begin{aligned}
&5{\cdot}921\mathrm{D}^4 + 65{\cdot}93\mathbf{k}\mathrm{YD}^3 + (0{\cdot}3428 + 7{\cdot}820\mathbf{k}\mathrm{Y})\,\mathrm{D}^2 + 2{\cdot}488\mathbf{k}\mathrm{YD} \\
&\qquad\qquad + 0{\cdot}002812 + 0{\cdot}3754\mathbf{k}\mathrm{Y} = 0.
\end{aligned} \tag{41}$$

The real and imaginary portions of the roots of the latter equation have been plotted in fig. 6 in order to indicate the degree of initial stability possessed by the arrangement. When $\mathrm{Y} = 0$ the two truck-elements are independent and the roots are given by equation 28. We are thus able to associate one pair of curves ($\xi\eta$) with the main frame, and the other pair ($\xi'\eta'$) with the bogie. The periodic distances for the two elements approach equality for a time, but afterwards diverge, the motion associated with the main frame ultimately becoming aperiodic. For the first portion of the curves, and until $\mathbf{k}\mathrm{Y} = 0{\cdot}0211$—a figure which corresponds with a centring force (2Y) of 2·86 metric tons per cm. displacement—the motion is stable, but beyond this limit it is unstable and the instability is in the motion associated with the bogie. Fig. 7 shows the limiting speed of stable running, deduced from the characteristic equation 40. We infer that, beyond the limits indicated in the figures, the bogie tends to lash the rails; but, being comparatively light, and connected elastically with the main mass of the locomotive, the impacts are unlikely to be a source of danger at ordinary speeds. We reserve further comment until we have considered motion in the reverse direction.

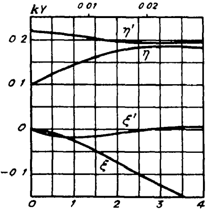


Fig. 6.—4-6-0 Locomotive. Roots of Characteristic Equation. Zero Speed.

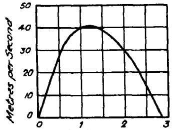


Fig. 7.—4-6-0 Locomotive. Speed-limit of Stable Running.

The characteristic equation for motion in the reverse direction, *i.e.*, with bogie in the rear, can be written down immediately from equation 40, the quantities **b**, **b′**, **c**, **c′**, *h*, *h′*, being reversed in sign. It is as follows:—

$$\begin{aligned}&0{\cdot}288\mathbf{a}^4\mathrm{D}^8 + 5{\cdot}595\mathbf{a}^3\mathrm{D}^7 + [26{\cdot}135\mathbf{a}^2 + (-0{\cdot}0240 + 5{\cdot}023\mathbf{kY})\,\mathbf{a}^3]\,\mathrm{D}^6\\ &+ [28{\cdot}00\mathbf{a} + (-0{\cdot}433 + 75{\cdot}74\mathbf{kY})\,\mathbf{a}^2]\,\mathrm{D}^5\\ &+ [5{\cdot}921 + (-1{\cdot}564 + 139{\cdot}6\mathbf{kY})\,\mathbf{a} + (0{\cdot}901 - 0{\cdot}480\mathbf{kY})\,\mathbf{a}^2]\,\mathrm{D}^4\\ &+ [65{\cdot}93\mathbf{kY} + (1{\cdot}2815 - 6{\cdot}856\mathbf{kY})\,\mathbf{a}]\,\mathrm{D}^3\\ &+ [0{\cdot}3428 - 7{\cdot}820\mathbf{kY} + (-0{\cdot}0751 + 2{\cdot}211\mathbf{kY})\,\mathbf{a}]\,\mathrm{D}^2\\ &+ 2{\cdot}488\mathbf{kY}\mathrm{D} + 0{\cdot}002812 - 0{\cdot}3754\mathbf{kY} = 0. \qquad (42)\end{aligned}$$

The region of stable motion now lies between $\mathbf{kY} = 0$ and $\mathbf{kY} = 0{\cdot}0075$, at which value the independent term in equation 42 passes through zero. The roots of the characteristic equation for zero speed are shown in fig. 8. The motion associated with the bogie preserves its periodic character throughout, without any great change in period; that associated with the main truck increases in period, and becomes aperiodic at about $kY = 0 \cdot 0066$, where the damping curve splits into two branches, one of which rises rapidly and becomes positive at about $kY = 0 \cdot 0075$. The curve of limiting speed for stable running is shown in fig. 9 for this case. We infer that beyond the limit indicated, the

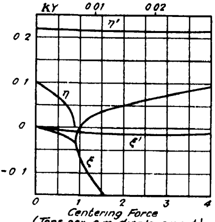


Fig. 8.—0-6-4 Locomotive. Roots of Characteristic Equation. Zero Speed

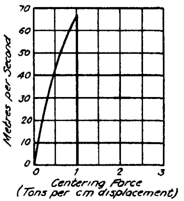


Fig. 9.—0-6-4 Locomotive. Speed-limit of Stable Running.

motion becomes unstable, not by increase in amplitude of a periodic factor but by a buckling of the wheel-base tending to cause derailment at a fore-wheel, and, moreover, that the impacts of the flanges on the rail, when the locomotive is running at speed, are backed by the mass of the main frame and its appurtenances, and are accordingly liable to constitute a source of danger.

Perhaps the chief conclusion to be drawn from the foregoing discussion is that, for the type of locomotive in question, the centring force for small displacements of the bogie should be moderate in magnitude, and no tendency to lock at the central position should be permitted. This is particularly necessary when the locomotive is intended to run at speed with bogie in the rear. It may be contended that the friction between the parts would be effective in opposing deviation of the trucks; and in so far as the question is that of running on a perfectly smooth track the contention may be justified. But where a local peculiarity in the track forces the trucks into an unusual configuration, the friction, by opposing their return, may constitute an element of danger. It is probable that a number of derailments of locomotives having rear bogies which have occurred, and for which no adequate explanation has been vouchsafed, are to be attributed to unstably running locomotives, having a natural tendency to derail, meeting with some concatenation of unfavourable track conditions.

We next consider locomotives in which the pivot is fixed in both main and auxiliary frames. Chief of these is the locomotive having, in addition to its main axles, a pony axle of the Bissel type; but more complex machines having the same characteristic are extant, such, for instance, as certain electric locomotives used on the Cascade Division of the Chicago Milwaukee and St. Paul Railroad, in which a six-wheel truck is so pivoted to the main frame. Since the pivot is fixed in both frames, $Y = \infty$, and the characteristic equation is formed from expressions 37 and 39 alone.

We take as an example the locomotive shown diagrammatically in fig. 10,

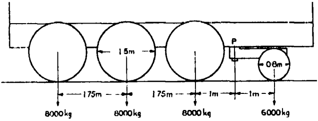


FIG. 10.—Bissel Locomotive (2–6–0 or 0–6–2).

which resembles that of fig 3, except in that the axle bearing the smaller wheels is carried in an auxiliary frame which is pivoted at P to the main frame. The mass of the auxiliary frame is small and is here neglected, so that $\mathbf{a}' = 0$ The constants involved when small wheels are ahead are as follows:—

$f = 1 \cdot 96 \times 10^{6}$ $\mathbf{f} = 5 \cdot 88 \times 10^{6}$ $\mathbf{k} = 1 \cdot 7 \times 10^{-7}$

$f' = 1 \cdot 24 \times 10^{6}$ $= \mathbf{f}'$ $\mathbf{k}' = 8 \cdot 07 \times 10^{-7}$

$\mathbf{c} = -2 \cdot 75$ $\mathbf{c}' = 1$

$\mathbf{s} = 0 \cdot 05$ $\mathbf{s}' = 0 \cdot 0938$

$\mathbf{b} - \mathbf{c} = 2 \cdot 162$ $\mathbf{b}' - \mathbf{c}' = -0 \cdot 906$

$\mathbf{d}^2 = 10 \cdot 167$ $\mathbf{d}'^2 = 1 \cdot 5625$

$\mathbf{d}^2 - \mathbf{c}^2 = 2 \cdot 604$ $\mathbf{d}'^2 - \mathbf{c}'^2 = 0 \cdot 5625$

$\mathbf{a} = 5 \cdot 2 \times 10^{-4}\, V^2$ $\mathbf{a}' = 0$

$h = -2$ $k^2 = 4$

$h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c} = 7 \cdot 167$ $h^2 + k^2 + \mathbf{d}'^2 - 2h\mathbf{c}' = 13 \cdot 56.$

The characteristic equation of the locomotive is therefore

$$
\begin{aligned}
&50\cdot44\mathbf{a}^2\mathrm{D}^5 + [98\cdot0\mathbf{a} + (-29\cdot25 + 153\cdot1\mathbf{k}\mathrm{G})\,\mathbf{a}^2]\,\mathrm{D}^4 \\
&+ [42\cdot56 + (-43\cdot42 + 383\cdot7\mathbf{k}\mathrm{G})\,\mathbf{a}]\,\mathrm{D}^3 \\
&+ [-16\cdot55 + 239\cdot8\mathbf{k}\mathrm{G} + (-3\cdot937 + 5\cdot355\mathbf{k}\mathrm{G})\,\mathbf{a}]\,\mathrm{D}^2 \\
&+ [2\cdot251 + 1\cdot325\mathbf{k}\mathrm{G}]\,\mathrm{D} + 0\cdot0509 + 3\cdot235\mathbf{k}\mathrm{G} = 0. \qquad (43)
\end{aligned}
$$

For slow motion, the limiting condition of stability requires $\mathbf{k}\mathrm{G} > 0\cdot096$, or a minimum aligning couple (2G) of about 11·3 metric tons per cm. movement at a radius of 1 metre. This is impracticably large as a proportional couple throughout the whole range of movement necessary. Fig. 11 shows the roots

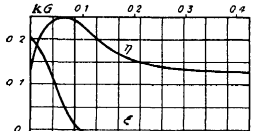


FIG. 11.—2-6-0 Locomotive Roots of Characteristic Equation. Zero Speed

of equation 43 when $\mathbf{a} = 0$. It is not now possible to associate the curves with particular trucks, since the co-ordinates are no longer independent. The curves indicate a simple motion of deviation, always stable, and periodic motion of varying frequency, initially unstable, but becoming and remaining stable; its ultimate value, when $\mathrm{G} = \infty$, corresponds with the case of fig. 3. Fig. 12 shows the limiting speed for stable running as a function of the aligning couple.

For motion in the reverse direction, *i.e.*, with trailing axle, the characteristic equation becomes—

$$
\begin{aligned}
&50\cdot44\mathbf{a}^2\mathrm{D}^5 + [98\cdot0a + (29\cdot25 + 153\cdot1\mathbf{k}\mathrm{G})\,\mathbf{a}^2]\,\mathrm{D}^4 \\
&+ [42\cdot56 + (43\cdot42 + 383\cdot7\mathbf{k}\mathrm{G})\,\mathbf{a}]\,\mathrm{D}^3 \\
&+ [16\cdot55 + 239\cdot8\mathbf{k}\mathrm{G} - (3\cdot937 + 5\cdot355\mathbf{k}\mathrm{G})\,\mathbf{a}]\,\mathrm{D}^2 \\
&+ [2\cdot251 - 1\cdot325\mathbf{k}\mathrm{G}]\,\mathrm{D} - 0\cdot0509 + 3\cdot235\mathbf{k}\mathrm{G} = 0. \qquad (44)
\end{aligned}
$$

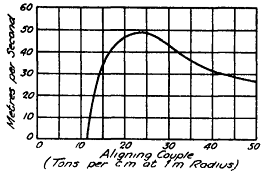


FIG 12 —2-6-0 Locomotive. Speed-limit of Stable Running.

Fig 13 shows the roots of this equation when a = 0. The fact that stability does not commence until kG = 0·0509/3·235, is due to the negative sign of the figure 0·0509 in the equation; this figure, however, arises as the difference of two nearly equal quantities and its sign appears to be accidental rather than characteristic of the type. The upper limit of stability is at kG = 1·29, or 2G = 152 tons per cm. at 1 metre radius. At G = 0 the motion has a periodic component, which quickly becomes aperiodic and remains at this until a little after stability has been established, when it again becomes periodic and remains so until G = ∞. Fig. 14 gives the limiting speed for stable running for this case.

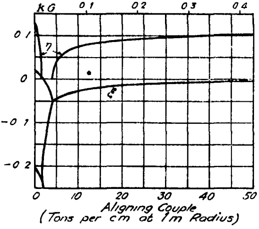


FIG 13 —0-6-2 Locomotive Roots of Characteristic Equation. Zero Speed

Locomotives of the type now under notice, when run with Bissel ahead, require so strong an aligning couple that it is not unreasonable practice which locks the truck somewhat strongly at its central position. With too small a couple, nosing is likely to become manifest when running at speed. The couple which yields the greatest stability is seen from fig. 12 to be a little greater than corresponds with the lower limit of stability, and this might be obtained by the use of suitable compound springs or equivalent devices. The permissible range of this couple is, however, on account of its high value, so small, that it must usually be surpassed when running on curves and even on straight track the possible deviation of the truck may exceed the permissible range, and become manifest in unstable running. Locomotives of the type are, therefore, rightly considered unsuitable for high speed working. Running the locomotive in the reverse direction, however, the trailing Bissel has a stabilizing effect for a large and useful range of values of aligning couple.

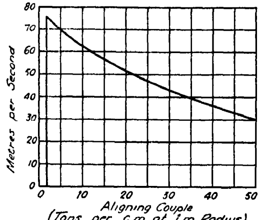


FIG. 14.—0-6-2 Locomotive. Speed-limit of Stable Running.

The provision of an aligning couple on an auxiliary truck, in addition to a centring force on its pivot, is not common. It was proposed by the author for the rear bogies of high speed locomotives,* under the inspiration of some dim and partial apprehension of the performance. It is employed on the bogies of certain Brown-Boveri-Winterthur electric locomotives recently put into service on the Paris-Orleans Railway.† The author also proposed‡ changing the pivotal centres of the auxiliary trucks in symmetrical locomotives of 4-6-4, 4-6-6-4, and similar types, so that the front truck is used as a simple bogie, having such characteristics as shown in figs. 6 and 7, whilst the rear truck trails, and has general characteristics such as shown in figs. 13 and 14. It was the intention that the locomotive should run equally in either direction, the change of pivots being brought about by the reversing gear as part of the control system. Very stable running would be expected of such a locomotive. Arrangements for changing the pivotal centres were provided, as an experimental measure, on one of the Brown-Boveri-Winterthur locomotives referred to above, but the results of tests have not yet been disclosed.

---

* See British Patents 128106, 155038.

† Parodi, "Électrification partielle du réseau de la Compagnie d'Orléans," 'Revue Générale des Chemins de Fer,' 1927, 46e année, 2me semestre, p. 23, fig. 2.

‡ See British Patent 163185.

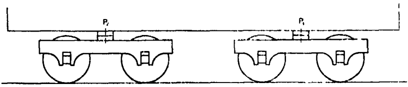


FIG. 15.—Double Bogie Locomotive (4–0–4).

A common type of locomotive of two axle groups is the double bogie locomotive shown diagrammatically in fig 15; and the bogie coach, whether motor or trailer, may be included in the same category for the present purpose. This, since it involves three parts, the two bogies and the body, leads in general to a characteristic equation of the twelfth degree,* which, however, in the case of a symmetrical locomotive, is readily split into two of the sixth degree, as will be shown. We consider the symmetrical case only.

We take the centre of each bogie as reference point for its own motion, distinguishing between them by suffixes 1 and 2, and the centre of the body as reference point for its motion, using co-ordinate without distinguishing marks. Let M be the mass of the body and I its moment of inertia about the vertical axis through its centre. Let $2s$ be the distance between bogie centres. For the bogies, with the usual notation, $\mathbf{b} = 0$, $\mathbf{c} = 0$, $h = 0$, also

$\mathbf{a}_1 = \mathbf{a}_2 (= \mathbf{a})$, $\mathbf{d}_1 = \mathbf{d}_2 (= \mathbf{d})$, $\mathbf{s}_1 = \mathbf{s}_2 (= \mathbf{s})$, $\mathbf{k}_1 = \mathbf{k}_2 (= \mathbf{k})$, $k_1 = k_2 (= k)$.

From equations 23 and 24,

$$(\mathbf{a}\mathrm{D}^2 + \mathrm{D})\, y_1 - \psi_1 = \tfrac{1}{2}\mathbf{k}\mathrm{Y}_1', \tag{45}$$

$$(\mathbf{a}\mathrm{D}^2 + \mathrm{D})\, y_2 - \psi_2 = \tfrac{1}{2}\mathbf{k}\mathrm{Y}_2', \tag{46}$$

$$\mathbf{s}y_1 + (\mathbf{a}k^2\mathrm{D}^2 + \mathbf{d}^2\mathrm{D})\, \psi_1 = 0, \tag{47}$$

$$\mathbf{s}y_2 + (\mathbf{a}k^2\mathrm{D}^2 + \mathbf{d}^2\mathrm{D})\, \psi_2 = 0; \tag{48}$$

also

$$\mathrm{MV}^2\mathrm{D}^2 y = -(\mathrm{Y}_1' + \mathrm{Y}_2'), \tag{49}$$

$$\mathrm{IV}^2\mathrm{D}^2 \psi = -s(\mathrm{Y}_1' - \mathrm{Y}_2'), \tag{50}$$

and we may write

$$\mathrm{Y}_1' = -2\mathrm{Y}(y_1 - y - s\psi), \tag{51}$$

$$\mathrm{Y}_2' = -2\mathrm{Y}(y_2 - y + s\psi). \tag{52}$$

---

* See § 9 below

Adding equations 45 and 46, 47 and 48, and substituting in the results, and in equation 49, for Y_1' Y_2' gives

$$(\mathbf{a}\mathrm{D}^2 + \mathrm{D} + \mathbf{k}\mathrm{Y})(y_1 + y_2) - (\psi_1 + \psi_2) - 2\mathbf{k}\mathrm{Y}y = 0, \tag{53}$$

$$\mathbf{s}(y_1 + y_2) + (\mathbf{a}k^2\mathrm{D}^2 + \mathbf{d}^2\mathrm{D})(\psi_1 + \psi_2) = 0, \tag{54}$$

$$2\mathrm{Y}(y_1 + y_2) - (\mathrm{M}\mathrm{V}^2\mathrm{D}^2 + 4\mathrm{Y})y = 0; \tag{55}$$

writing $\alpha\mathbf{a} = \frac{1}{4}\mathbf{k}\mathrm{M}\mathrm{V}^2$ these equations yield

$$\alpha[\mathbf{a}^3k^2\mathrm{D}^6 + \mathbf{a}^2(k^2 + \mathbf{d}^2)\mathrm{D}^5 + \mathbf{a}\mathbf{d}^2\mathrm{D}^4 + \mathbf{a}\mathbf{s}\mathrm{D}^2]$$
$$+ \mathbf{k}\mathrm{Y}[(1 + \alpha)\mathbf{a}^2k^2\mathrm{D}^4 + \mathbf{a}(k^2 + \mathbf{d}^2 + \alpha\mathbf{d}^2)\mathrm{D}^3 + \mathbf{d}^2\mathrm{D}^2 + \mathbf{s}] = 0. \tag{56}$$

This is a partial characteristic equation of the locomotive, and, since there is no term in D, it represents an unstable motion. In a similar manner, by taking the differences of equations 45 and 46, and of 47 and 48, substituting for Y_1' and Y_2' in these, and in equation 50, the remaining components of the motion are found. The resulting equation is similar to equation 56, with β written in place of α, where $\beta\mathbf{a} = \frac{1}{4}\mathbf{k}\mathrm{I}\mathrm{V}^2/s^2$, and therefore again represents an unstable motion. The destructive effective of the instability is, however, limited on account of the comparatively small mass of the trucks.

9. *Locomotives of more than Two Axle-Groups*

The general locomotive-unit may be said to comprise three axle-groups, a main group, with auxiliary groups fore and aft of it. The terms "main" and "auxiliary" must be considered as having reference to the frame through which draught is effected rather than to primary function: for, although in the steam locomotive the main axle-group comprises the driving axles, the auxiliary groups being used for steering and weight-distributing purposes, in the electric locomotive there is no such functional distinction, and auxiliary groups may consist partly or wholly of driving axles, indeed, in the double-bogie locomotive (4-0-4) the main group is lacking. A complete locomotive may consist of two or more units, but in such locomotives some of the intermediate auxiliary axle-groups are often omitted. Thus the locomotive usually represented by the symbol 4-6-6-4 is a two-unit locomotive lacking in two auxiliary axle-groups, and might with propriety, certainly with increased lucidity, be represented by the symbol 4-6-0 + 0-6-4. When two units are connected by means of a draw-bar, they are, for first order deviations, independent of one another; but when the connection takes the form of a Mallet hinge the locomotive must, for the present purpose, be considered as a whole. The hinge coupling is, however, unusual for locomotives intended to run at high speeds.

Each axle group in general adds four to the degree of the characteristic equation, but each fixed pivot or hinge joint reduces the degree again by two. Thus the locomotive 4–6–6–4, if the units are connected by a draw-bar, has a characteristic equation of the sixteenth degree, which, if the units are similar and the action of intermediate side buffers is neglected, can be divided into two equations, each of the eighth degree, related to one another in a manner similar to equations 40 and 42. If, however, the units are connected by a hinge, the characteristic equation of the locomotive is of the fourteenth degree. It is apparent, therefore, that the characteristic equation increases rapidly in length and complexity as the number of axle-groups is increased. No difficulty arises in forming the equations of deviation for any locomotive, but the deduction therefrom of the characteristic equation involves considerable labour when the number of axle-groups exceeds two.

It is not the author's purpose, in the present communication, to discuss the calculation of more complex cases further. Sufficient has been said to demonstrate the proposed treatment of the problem of stability, and the methods for deducing the characteristic equation of the locomotive, on which stability depends. The general effect of certain usual wheel arrangements has been discussed with the aid of particular examples; and from such partial studies the general performance in more complex cases may be inferred. Certain of these cases are perhaps worthy of a more complete treatment which may be accorded them hereafter as occasion demands.

APPENDIX—NOTATION.

The following notation is used in the paper—dashes and suffixes being used to distinguish symbols of a similar nature :—

a, radius of wheel, in millimetres in equations 1 and 2, elsewhere in metres.

b, a half the distance between the treads of the two wheels of a pair—a little more than a half the gauge of track; taken as 0·75 m. for standard gauge.

x, y, co-ordinates of a selected reference point on the longitudinal axis of symmetry of the locomotive or truck plan, referred to fixed axes, x is measured in the direction of the track, and y from the centre line of the track.

e, distance of an axle ahead of the selected reference point (xy).

h, distance of the centre of inertia of a truck ahead of the selected reference point (xy).

s, distance of a reference point for outside force system ahead of the selected reference point (xy).

f, creepage coefficient, defined and explained in section 3.

k, radius of gyration of a truck about a vertical axis through its centre of inertia.

m, M, masses, generally of truck or other assembly of parts associated in movement. [N B.—Since forces are herein taken in gravitational units (kg.), inertial masses must be divided by g, see **a** below]

w, weight borne by a wheel (kg.).

mk^2, I, moment of inertia of truck, or other assembly of parts.

X, component force on truck in direction of track, also used for periodic distance.

Y, component force on truck transverse to track.

G, couple acting on truck about a vertical axis.

V, velocity of progression of locomotive, metres per sec.

α, β, ratios, explained in section 8.

γ, creepage, defined in section 3.

λ, conicity of wheels, generally 1/20.

ψ, angle between axis of wheel-base of locomotive or truck and line of track.

In the following, the summation is taken over the axles of an associated group :—

$$\mathbf{f} = \Sigma(f), \quad \mathbf{k} = 1/\Sigma(f), \quad \mathbf{a} = mV^2/2g\Sigma(f), \quad \mathbf{b} = \Sigma\left(f\lambda\frac{b}{a}c\right)\Big/\Sigma(f).$$

$$\mathbf{c} = \Sigma(fc)/\Sigma(f), \quad \mathbf{d}^2 = b^2 + \Sigma(fc^2)/\Sigma(f), \quad \mathbf{s} = \Sigma\left(f\lambda\frac{b}{a}\right)\Big/\Sigma(f)$$

Here **a**, **b**, **c** and **d** are linear magnitudes, but **s** is a number, and **f** a force.

When, as is usually the case in practice, the wheels of an axle-group are of equal radius and carry equal loads, f is the same for all axles and the expressions given above are simplified. Thus, if n is the number of axles in the group, $\mathbf{f} = nf$, $\mathbf{c} = \frac{1}{n}\Sigma(c)$, $\mathbf{s} = \lambda\frac{b}{a}$, and so on.

We may note the following expressions of frequent occurrence :—

$$\mathbf{b} - \mathbf{sc} = \lambda b\Sigma\left[f_r f_s (c_r - c_s)\left(\frac{1}{a_r} - \frac{1}{a_s}\right)\right]\Big/[\Sigma(f)]^2.$$

This, a function of the differences of the c's, is independent of the selected reference point. It is zero when the wheels of the group are of equal radius, whether or not they are equally loaded.

$$\mathbf{d}^2 - \mathbf{c}^2 = b^2 + \Sigma\,[f_r f_s (c_r - c_s)^2]/[\Sigma\,(f)]^2,$$

$$h^2 + k^2 + \mathbf{d}^2 - 2h\mathbf{c} = k^2 + b^2 + \Sigma\,[f\,(h - c)^2]/\Sigma\,(f),$$

$$h^2 + k^2 + \mathbf{d}'^2 - 2h\mathbf{c}' = k^2 + b^2 + \Sigma\,[f'\,(h - c')^2]/\Sigma\,(f').$$

These expressions are also independent of the choice of reference point.

The Spacial Distribution of Photoelectrons produced by X-Rays.

By E. J. Williams, J. M. Nuttall, and H. S. Barlow.

(Communicated by W. L. Bragg, F.R.S.—Received August 7, 1928.)

Introduction.

Investigations on the initial directions of emission of photoelectrons produced by X-rays show that the distribution depends on the direction of propagation of the rays, whilst in the case of a polarised beam it depends also on the direction of the electric vector of the incident radiation. The distribution of the initial directions of the photoelectrons relative to the direction of the incident X-ray beam is usually called the "longitudinal distribution," whilst that relative to the plane containing the electric vector and the rays, in the case of a polarised beam, is termed the "lateral distribution."

The present paper will deal only with the longitudinal distribution of the photoelectrons. In the first place a survey of the experimental situation will be given, from which it will be seen that the results of previous observers are not in a satisfactory state of agreement. Secondly, it is the purpose of this paper to give an account of the authors' own experiments on the longitudinal distribution in which it is hoped that some of the possible errors giving rise to the discordant results of previous investigators have been eliminated. The main interest of the present work lies in the results obtained for the asymmetry, which show that the average forward momentum of a photoelectron is nearly

40 per cent. greater than that of an absorbed quantum. This result of our experiments was first announced in a thesis published in April, 1927.*

Theory and Previous Experimental Results.

The experimental observations of all investigators on the longitudinal distribution of the photoelectrons are in agreement in the following two important respects :—

(1) There is an emission of photoelectrons over a wide range of angles, which may conveniently be called the " spread " or " dispersion " of the photoelectrons.

(2) There is a much greater emission of photoelectrons in the forward direction of the rays than in the backward direction : this is generally termed the longitudinal " asymmetry." These two features of the distribution may to a considerable extent be considered separately. Various theories have been advanced to account for the dispersion of the photoelectrons, one of the most recent being that given by Auger and Perrin.† According to the latter, if we disregard the second order effect of the asymmetry, the dispersion is independent of the atom from which the photoelectron is ejected and of the wave-length of the incident X-rays, and depends only on the direction of the electric force in the rays, the probability of emission in any direction being proportional to the mean square of the component of the electric force in that direction. This law, usually called the " $\cos^2\theta$ " law, was first enunciated by Auger and Perrin and subsequently deduced by Wentzel,‡ using the ideas of wave mechanics

The explanation of the longitudinal asymmetry is attributed to the effect of the magnetic force in the incident wave, or in terms of the light-quantum theory, it is assumed that the incident X-ray quantum communicates its momentum, $h\nu/c$, to the ejected photoelectron, the forward momentum of the photoelectrons therefore being equal to that of the absorbed radiation. In support of this, experimental observations indicate that the asymmetry of the photoelectron distribution becomes more pronounced as the frequency of the incident X-rays increases.

---

* E. J. Williams, Thesis on 'The initial directions of Photoelectrons,' presented to the Commissioners for the Exhibition of 1851 in April, 1927. [Mr. Nuttall and Mr. Barlow kindly permitted me to give in this thesis an account of the results we had obtained up to that time.—E.J.W.]

† 'C.R.,' vol. 150, p. 1742 (1925).

‡ 'Z. Physik,' vol. 40, p. 578 (1927).

The application of the above mentioned theories to previous experimental results has been recently examined and discussed by one of us in a letter to 'Nature.'* It is pointed out there, that whilst the experimental results for the dispersion are in fair agreement and approximately satisfy the $\cos^2 \theta$ law, the results for the longitudinal asymmetry are not in agreement, and with the exception of those of Auger, demand an asymmetry greater than that expected from the supposition that the average forward momentum of a photoelectron is equal to the momentum of an absorbed quantum.

In estimating the magnitude of the asymmetry in the longitudinal distribution of the photoelectrons produced by X-rays the quantities most frequently measured by observers are (1) the ratio, ρ, of the number of photoelectrons emitted in a forward to those emitted in a backward direction, and (2) the value of the bipartition angle, ϕ_b, which is the semi-angle at the apex of a cone having the X-ray beam as its axis and dividing the photoelectron distribution into two equal parts. A third quantity which may also be determined and which is a significant measure of the asymmetry is the average value, $\overline{\cos \phi}$, of the cosine of the angle which the initial direction of the photoelectron makes with the X-ray beam. As hitherto there has been no uniform practice adopted by observers in stating the magnitude of the asymmetry it seems desirable to introduce a definite asymmetry factor. This will enable us to correlate the results of previous observers and later to compare them with our own. The asymmetry factor, σ, which we shall use in this paper is defined in such a way that its value is unity, if the average forward momentum of the photoelectron, μ, is equal to $h\nu/c$ where ν is the frequency of the incident radiation, *i.e.*, $\mu = \sigma(h\nu/c)$. σ is defined according to the following scheme. Consider a system of momentum vectors, T_e, with a direction distribution such that they obey the $\cos^2 \theta$ law and let a constant momentum, T_m, in the forward direction be added to each, the magnitude of T_e being such that the resultant velocities corresponding to the momenta $(T_e + T_m)$ satisfy Einstein's law for the energy of a photoelectron. It is assumed that the value of T_m is appropriately chosen so that the resultant distribution obtained, represents the observed distribution of emission of the photoelectrons. Clearly the magnitude of the longitudinal asymmetry will increase with increasing values of T_m. Now if the added vector, T_m, is equal to $h\nu/c$, it can be shown that the *average* forward momentum of a photoelectron, μ, is equal to $4/5 \,.\, h\nu/c$, and hence if the average forward momentum of the photoelectrons is equal to $h\nu/c$ then T_m must equal $5/4 \,.\, h\nu/c$. We accordingly define σ as $T_m \div 5/4 \,.\, h\nu/c$, the factor 5/4 being introduced simply to make the value of σ unity when the average value of the forward momentum of a photoelectron μ is equal to the momentum $h\nu/c$ of the incident quantum. It can then be shown that the quantities usually measured in determining the longitudinal asymmetry, viz., $\rho \,.\, \overline{\cos\phi}$ and ϕ_b are expressed in terms of the asymmetry factor, σ, by the following relations—

---

* Williams, 'Nature,' vol. 121, p. 134 (1928).

$$(\rho - 1)/(\rho + 1) = 15/8 \,.\, k\sigma \text{ (approx.)} \tag{1}$$

$$\overline{\cos\phi} = k\sigma \tag{2}$$

$$\cos\phi_b = 5/4 \,.\, k\sigma \tag{3}$$

where $k = \sqrt{h/2mc\lambda}$; λ is the wave-length of the incident X-rays and the binding energy of the photoelectrons is neglected for simplicity. The values of σ have been calculated from the results of previous observers using the above expressions, and are given in Table I.

Table I.

Gas.	λ in Å	Number of photo-electrons examined	σ calculated from			Observers.
			$\overline{\cos\phi}$	ρ	ϕ_b	
Air	0·37 to 1·0	1000	--	1·5	--	C. T. R. Wilson*
Air	0·71	250	—	2·1	—	O. K. Defoe†
O_2 + argon	0·71	150	2·0	2·1	2·0	D. H. Loughridge‡
Argon	0·13	300		1·4	1·3	F. Kirchner§
Air	0·61	200		1·0	0·83	P. Auger‖
Air	0·81	200	-	0·9	0·83	,,
Argon	0·15	200		0·9	0·86	,,

* 'Roy. Soc. Proc.,' A, vol. 104, p. 1 (1923).
† 'Phil. Mag.,' vol. 49, p. 817 (1925).
‡ 'Phys. Rev.,' vol. 30, p. 488 (1927).
§ 'Ann. Physik,' vol. 84, p. 899 (1927).
‖ 'J. de Phys.,' vol. 8, p. 85 (1927).

The table shows that previous results correspond to values of σ varying from about 1 to 2, and, with the exception of those of Auger, they indicate that the average value of the forward momentum of a photoelectron is greater than the momentum of an absorbed quantum. The latter interpretation of the absolute value of σ in the cases where it has been deduced from observations on the quantities ρ and ϕ_b is, of course, dependent on the scheme used in defining σ, but the relative values of σ are independent of this scheme and clearly show that the results of all other observers correspond to an asymmetry in the distribution greater than that observed by Auger.

Although there is not a very satisfactory agreement between the results of different observers, yet on the whole they do seem to indicate that the longitudinal asymmetry cannot be quantitatively explained by assuming that the momentum of the absorbed radiation is transferred to the ejected photoelectron.

In attempting to explain the variations in the experimental results quoted in Table I it is important to note that the Wilson cloud method of observing the photoelectrons was used in all the investigations. As the same method has also been used in the authors' own experiments, a discussion of the possible errors arising from the application of this method (as well as from other sources), and the precautions adopted to eliminate them in the present investigation, will now be given.

Discussion of Possible Errors in the Observation and Measurements of the Photoelectron Tracks.

(1) ***Wave-length of the Radiation Producing the Photoelectrons.***– Since the longitudinal asymmetry varies with the wave-length of the X-rays producing the photoelectrons, it is important that homogeneous rays be used. The number of photoelectrons produced in the cloud chamber by radiation of wave-length between λ and $\lambda + d\lambda$, and possessing a given amount of energy is approximately proportional to λ^4, and the effective wave-length of a beam of heterogeneous rays may therefore be very different from the wave-length which represents the maximum concentration of energy. If an inadequate correction is applied for this effect, the observed asymmetry will be too small since the asymmetry decreases with increasing wave-length. In the present experiments homogeneous X-rays produced by reflection from a crystal were used and the range of wave-lengths in the reflected rays was only about 2 or 3 per cent. of the mean value There can, therefore, be no question of errors arising from this source. In his observations C. T R Wilson deduced the wave-length of the X-rays from the range of the photoelectrons produced. All other observers have used radiation direct from an X-ray tube, passed through suitable filtering screens, the mean wave-length being estimated from the nature of the primary radiation and the filters used.

2. ***Method of Examining the Photographs.***—To deduce the direction of a photoelectron track in space it is necessary to take photographs from two different points of view. In the present work stereoscopic photography was used, and by virtue of the method employed in examining the photographs there is no possibility of errors arising from an inaccurate knowledge of the stereoscopic angle, the magnification, etc., because such a knowledge is not

required The method consists in replacing the photographic negatives in the cameras and reconstructing the tracks in front. It was first used by two of the authors in a previous investigation* on the ranges of secondary β-rays produced by X-rays. We hope later to give a fuller account of the method elsewhere. The beam of X-rays used to produce the photoelectrons was in the form of a thin sheet so that the direction of the X-rays with respect to the camera could be determined by observing the positions of the origins of different photoelectrons.

3. *The Effect of "Choosing" Tracks on which Measurement is made.*—In describing the procedure adopted in examining the photographs, some observers state that they "choose" for measurement, those photoelectron tracks which can be measured most accurately. Whilst this undoubtedly leads to greater accuracy in measurements on individual tracks, it may easily have an adverse effect on the accuracy of the statistical results This follows because tracks which make small angles with the axes of the cameras appear more difficult to measure than others, and are therefore systematically omitted if choosing of the above-mentioned type is adopted This effect, more than any other, possibly accounts for the variations between the results of different observers. The way we tried to eliminate it in the present measurements is described later.

4. *Statistical Errors arising from the Inaccurate Measurement of the Directions of Individual Photoelectrons.*—The effect of random errors in the measurement of the initial direction of individual photoelectrons is systematic in diminishing the asymmetry and increasing the dispersion. As far as such errors are concerned the observed asymmetry must therefore be regarded as the minimum whilst the dispersion observed must be the maximum effect Auger and Loughridge made observations on photoelectrons produced in an atmosphere of hydrogen containing a small proportion of a heavier gas. Under such conditions the photoelectrons travel a considerable distance before they are deflected and the initial direction can be measured with greater accuracy. As a hydrogen atmosphere was not used in the author's experiments, the conditions did not permit quite as high accuracy in the measurements on individual tracks as is possible in the investigations above mentioned. However, in virtue of the nature of the statistical effect of these errors as indicated above, the main significance of our results is not impaired by this fact. We will now proceed to consider the experimental results.

---

* Nuttall and Williams, 'Phil. Mag.,' vol 2, p 1109 (1926).

Experimental Results.

The experimental method of obtaining the photographs of the photoelectron tracks has already been described in previous papers.* In all cases homogeneous X-rays were used of wave-lengths 0·545 Å., 0·614 Å. and 0·709 Å., oxygen and nitrogen being the gases introduced into the cloud chamber.

1. *The Ratio,* ρ, *of Forward to Backward Emission of Photoelectrons.*—In order to determine the ratio, ρ, it is, of course, only necessary to observe whether the angle, ϕ, between the initial direction of a photoelectron and the direction of the X-rays, is greater or less than 90°. For most photoelectron tracks this can be ascertained by mere inspection of the photographs, and it is possible to make observations on a large number of photoelectrons in a comparatively short time. In view of this a special study of the value of ρ was made. In this study, observations were made on *every* track on *every* photographic plate examined, and the results are essentially free from systematic choosing. The results are given in Table II. The gas traversed and the wave-length of the X-rays used are given in the first and second columns. The actual numbers of photoelectrons involved appear in the third, fourth and fifth columns, and the corresponding values of ρ in the sixth. The asymmetry is measured by the extent that ρ differs from unity, and an asymmetry factor, Sρ, may be conveniently defined as (ρ − 1)/(ρ + 1). The values of this quantity are given in the last column of the table.

Table II.

Gas.	λ in Å.	Total number of photo-electrons examined	Number emitted in forward direction N_F	Number emitted in backward direction N_B	$\rho = \frac{N_F}{N_B}$	$S_\rho = \frac{\rho - 1}{\rho + 1}$
Nitrogen	0·614	532	361	171	2·11	0·357
Nitrogen	0·545	148	99	49	2·02	0·337
Oxygen	0·709	182	121	64	1·96	0·324
Oxygen	0·614	613	416	197	2·11	0·357
Oxygen	0·545	359	246	113	2·18	0·373
Mean, all cases	0·60	1837	1243	594	2·09	0·353

The range of wave-lengths employed in these experiments is not sufficient to show a very marked variation in ρ, although they do give a general indication that ρ increases for decreasing values of λ. On the whole, however, the variation of ρ with the nature of the gas and with the wave-length of the rays is of about the same order of magnitude as the statistical fluctuations. Under these circumstances and considering that the wave-length for each case is accurately known as well as the exact contribution to the total number of photoelectrons, the most accurate result is obtained by averaging the results as indicated in the last line of the above table. The effective wave-length for all cases is 0·60 Å., and the mean value of ρ, corresponding to this wave-length, for oxygen and nitrogen, is 2·09, the corresponding value of S_ρ being 0·353 This result is based on observations of more than 1800 photoelectrons and the error arising from the finiteness of the number of tracks examined is such that there is an even chance of an error less than 0·04 in the value of ρ and a four to one chance of an error less than 0·07. Auger made observations on the photoelectron tracks in air using X-rays of wave-lengths corresponding to 20,000 and 15,000 electron volts, *i.e.*, of wave-lengths 0·61 Å. and 0·82 Å respectively. The conditions are therefore approximately the same as those of the present experiments. Auger, however, finds values of ρ in these cases equal to 1·77 and 1·46 respectively. Our results, therefore, do not substantiate the low values of the asymmetry observed by Auger. The results of Defoe and Loughridge for X-rays of wave-length 0·71 Å. traversing air correspond, on the other hand, to a value of ρ equal to 2·9 Our value for ρ is thus intermediate between those obtained by Auger and by Defoe and Loughridge. According to Table I the value of σ deduced from C. T. R. Wilson's observations is 1 5, corresponding to a value of ρ equal to 2·2 for X-rays of wave-length 0·60 Å , which is in satisfactory agreement with the present authors' results.

---

* Nuttall and Williams, 'Phil. Mag.,' vol. 1, p. 1217 (1926).

2. *The Space Distribution of the Photoelectrons.*- In order to determine completely the longitudinal space distribution it is necessary to determine the actual values of the angle, ϕ, which each photoelectron track initially makes with the direction of the X-rays. The accuracy with which the angle ϕ can be measured is appreciably diminished in some cases by the general indistinctness of the photograph, and, in others, by the closeness of the origins of the photoelectrons and the consequent entanglement of the tracks. These factors, however, bear no relation to the initial direction of emission of the photoelectrons, and, therefore, in making the measurements we may omit such tracks without affecting the accuracy of the statistical results. It has been pointed out, however, that the initial direction of the tracks may also control the choosing in such a way that photoelectrons are systematically omitted if their initial directions of emission make small angles with the axis of the camera.

This possibility was not at first fully appreciated in the present investigation. By means of the following analysis it is possible to test if the choosing of tracks on which measurement is made was appreciably influenced by their direction of emission and to allow for it if such is the case. Suppose we have a number of planes all intersecting along the direction of propagation of the X-ray beam and making equal angles with each other, then provided we use unpolarised X-rays, the numbers of photoelectrons initially emitted between any pair of adjacent planes should be equal, *i.e.*, the "lateral" distribution is a random one. If however in the measurements on the directions of the photoelectrons those making small angles with the axis of the camera are systematically omitted, then the numbers of photoelectrons emitted between these equidistant planes will decrease as the angle between the planes and the axis of the camera diminishes. The test may easily be carried out since the measurements made in determining the longitudinal distribution are also sufficient to calculate the lateral distribution, *i.e.*, the distribution on a plane

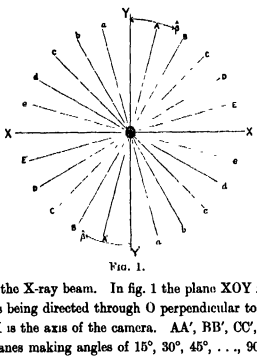


Fig. 1.

perpendicular to the X-ray beam. In fig. 1 the plane XOY represents such a plane, the X-rays being directed through O perpendicular to the plane of the paper, whilst OX is the axis of the camera. AA', BB', CC', . . , XX' represent traces of planes making angles of 15°, 30°, 45°, . . ., 90° with the plane YY'; *aa'*, *bb'*, . . . represent similar planes.

The numbers of photoelectrons, N_β, initially ejected between any two adjacent planes at intervals of 15° from 0° to 90° can be determined by observing the value of the angle, β, which each photoelectron initially makes with the plane YY'. The numbers are given in Table III.

Table III.

β.	0°–15°.	15°–30°.	30°–45°	45°–60°.	60°–75°.	75°–90°.
Number of photo-electrons N_β	186	193	187	170	135	100

For values of β up to 60° the value of N_β is constant within the probable statistical fluctuations.* There is, however, a considerable fall in the value of N_β in the region 60° to 75°, and the decrease in the last interval 75° to 90° is much greater than can be accounted for by any reasonable statistical fluctuation From the above-mentioned significance of the observed values of N_β it follows that photoelectrons for which the values of β are greater than 60° were to a certain extent systematically omitted, whilst up to 60° there is no evidence at all of such systematic choosing. The analysis thus reveals the existence of choosing of the kind suggested in the discussion of the possible sources of error. The same analysis may also be used to eliminate the detrimental effect of special choosing. The longitudinal distribution of photoelectrons produced by unpolarised X-rays is independent of the values of β which limit the region investigated, and hence by considering in the present experiments only those photoelectrons in the region for which the value of β lies between 0° and 60° we can obtain a distribution which is both representative and free from errors due to choosing. This procedure has been adopted. The value of ρ for the region $\beta = 0° - 60°$ is found to be 2·11, whilst its value obtained as the result of an entirely different examination of the photographs which was described in the previous section is 2·09. In the latter case *all* photoelectrons were included and the agreement between the two values of ρ substantiates the conclusion that the results for the region $\beta = 0 - 60°$ are free from errors due to choosing. The longitudinal distribution for this region will now be described in more detail.

3. *The Longitudinal Distribution of the Photoelectrons.* It has been previously pointed out that the determination of the complete longitudinal distribution involves the observation of the actual values of the angle ϕ, which the initial

---

* The X-radiation used to produce the photoelectrons was the characteristic radiation from a tube reflected from a crystal and owing to the latter process it was slightly polarised The angle of reflection was about 10°, and using Bubb's results for the lateral distribution for completely polarised rays it is found that the number of photoelectrons between 45° and 60° should be about 10 per cent. less than the average number for the four intervals from 0° to 60°. Thus, apart from statistical fluctuations, we can account for the slight drop in the number in the interval 45° to 60°.

portion of each photoelectron track makes with the direction of the X-rays. Observations of the angle ϕ were carried out for nearly 800 tracks representing photoelectrons emitted from oxygen and nitrogen by X-rays of wave-lengths 0·545 Å., 0·614 Å. and 0·709 Å. As emphasised in a previous discussion on the value of ρ, the variation in the longitudinal asymmetry is not very marked over this limited range and we have therefore combined the results in one table and graph, the results representing the photoelectric emission from oxygen and nitrogen due to radiation of mean effective wave-length 0·60 Å Table IV shows the number of photoelectrons, N_ϕ, emitted in intervals of 10° for values of ϕ from 0° to 180°, where ϕ is the angle between the initial direction of the photoelectrons and the direction of the X-rays The last column in the table gives the percentage number of photoelectrons emitted in these 10° angular intervals.

Table IV.

ϕ	Number of photo-electrons, N_ϕ	Percentage number of photo-electrons	ϕ	Number of photo-electrons, N_ϕ	Percentage number of photo-electrons
		Per cent			Per cent
0°–10°	4	0·5	90°–100°	72	9·2
10°–20°	18	2·3	100°–110°	68	8·7
20°–30°	25	3·2	110°–120°	53	6·9
30°–40°	44	5·6	120°–130°	35	4·5
40°–50°	61	7·8	130°–140°	8	1·1
50°–60°	78	10·0	140°–150°	11	1·4
60°–70°	88	11·3	150°–160°	9	1·1
70°–80°	103	13·2	160°–170°	1	0·1
80°–90°	101	12·8	170°–180°	0	0·0

The results of Table IV are represented graphically in fig. 2. The ordinates

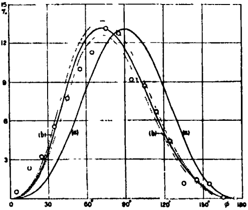


FIG. 2.—Percentage Number of Photoelectrons in 10° Angular Intervals (ϕ)

of the observed points (denoted by circles) represent the percentage number of photoelectrons for which ϕ lies within 5° of the abscissa of the point. The full curve marked (*a*) represents the theoretical distribution according to the $\cos^2\theta$ law assuming there is no asymmetry, *i.e.*, $\sigma = 0$. The magnitude of the observed dispersion is seen to be satisfactorily represented by this law, the intensity of the maximum concentration, for instance, being almost exactly the same for the two curves. The forces which give rise to the longitudinal asymmetry affect the dispersion slightly but it is only of the nature of a second order effect. The full curve marked (*b*) in fig. 1 represents the calculated distribution corresponding to an asymmetry factor $\sigma = 1 \cdot 37$, which is the mean value of σ calculated from our experimental values of $\overline{\cos\phi}$, ϕ_0 and ρ. The experimentally observed points are based on observations on about 800 tracks, and the dotted lines on each side of the full curve (*b*) represent the magnitude of the probable statistical fluctuations. These lines are drawn such that, owing to statistical fluctuations, half the observed points may lie outside the area enclosed by them. About half the observed points do actually lie within this area, and on the large angle side of the maximum the theoretical curve is probably the best representation of the experimental results. On the other side of the maximum the decrease in the observed number of photoelectrons is at first rather more marked, whilst at small angles the observed intensity is slightly greater than that represented by the theoretical curve. These minor differences should however not be stressed too much.

The main significance of the results and the points it is desired to emphasise are: -

(1) The dispersion experimentally observed is satisfactorily expressed by the $\cos^2\theta$ law; and

(2) The distribution observed indicates a "distortion" of the symmetrical $\cos^2\theta$ curve corresponding to a value of the asymmetry factor, σ, equal to about $1 \cdot 4$.

4. *The Average Forward Momentum of the Photoelectrons.*—If $\overline{\cos\phi}$ is the average value of the cosine of the angle, ϕ, which the initial direction of a photoelectron makes with the incident rays, then (taking into account the relativity correction and the binding energy of the electron), the average forward momentum, μ, of a photoelectron produced by a radiation of frequency, ν, is given by

$$\mu = m_0 c \cdot \overline{\cos\phi} \cdot \sqrt{(\alpha - \alpha_0)^2 + 2(\alpha - \alpha_0)} \qquad (4)$$

where $\alpha = h\nu/m_0c^2$ and $\alpha_0 = h\nu_0/m_0c^2$, $h\nu_0$ being the binding energy of the electron. From the definition of σ, we have $\mu = \sigma\,(h\nu/c)$, so that

$$\sigma\alpha = \overline{\cos\phi}\,\sqrt{(\alpha-\alpha_0)^2 + 2(\alpha-\alpha_0)}. \tag{5}$$

From the observed values of $\overline{\cos\phi}$, the known wave-length of the X-rays and the binding energy of the electrons, we may therefore deduce the value of σ, or the ratio of the average forward momentum of a photoelectron, to the momentum of an incident quantum. Table V gives the values of $\overline{\cos\phi}$ and σ for the various cases investigated.

Table V.

Gas.	λ in A.	Number of photoelectrons observed	$\overline{\cos\phi}$	σ
Nitrogen	0·614	200	0·189	1·32
Nitrogen	0·545	93	0·179	1·21
Oxygen	0·709	159	0·202	1·53
Oxygen	0·614	179	0·207	1·47
Oxygen	0·545	148	0·193	1·30
All cases	0·60 Å.	779	0·193	1·38

The variation of σ with the wave-length of the rays and the nature of the gas is within statistical fluctuations as in the case of the values of ρ. The mean result for all cases indicates that the average forward momentum of a photoelectron ejected from atoms of nitrogen and oxygen by radiation of wave-length approximately 0·60 Å is about 40 per cent. greater than the momentum of the incident quantum. The scheme we have used in defining σ is the same as that assumed in the theory of Perrin and Auger except they assume that the vector, T_m, which they call the *magnetic* impulse due to the radiation, is equal to the momentum $h\nu/c$ of an incident quantum. Since we have defined $T_m = 5/4 \,.\, h\nu/c$, it is seen that according to our results the magnetic impulse is about 1·7 times $h\nu/c$.*

---

* Equation (4) expresses the average momentum, μ, of a photoelectron *after* escaping from the field of the parent atom. The negative energy of an atomic electron is twice its ionising "potential," $h\nu_0$, so that the kinetic energy of a photoelectron before escaping from the atom is $h(\nu + \nu_0)$. According to Perrin and Auger ('Jour. de Phys.,' vol. 8, p. 104 (1927) the corresponding velocity must, in order to conform with wave mechanics, be directed along the outward drawn radius from the nucleus. The photoelectrons will therefore escape without being deflected, and the average forward momentum, μ', before escape is given by equation (4) with $(\alpha - \alpha_0)$ replaced by $(\alpha + \alpha_0)$. Owing to the smallness of α_0/α the numerical difference between μ and μ' is small. Actually the value of μ' deduced from the mean observed value of $\overline{\cos\phi}$ given in Table V is 1·40 and corresponds to a *magnetic impulse* of 1·75 times $h\nu/c$.

5. *The Relation between* ρ, $\overline{\cos\phi}$ *and* ϕ_b.—It was pointed out at the beginning of this paper that the magnitude of the longitudinal asymmetry may be estimated by measuring any one of the three quantities ρ, $\overline{\cos\phi}$ or ϕ_b. The determination of ρ and $\overline{\cos\phi}$ have already been discussed. The bipartition angle, ϕ_b, may readily be evaluated from the results given in Table IV, *i.e.*, by using the same observations from which values of ρ and $\overline{\cos\phi}$ have already been deduced. The final values of the quantities ρ, $\overline{\cos\phi}$ and ϕ_b are shown in Table VI

Table VI.

Total number of photoelectron tracks examined = 779.

ρ.	$S_\rho \dfrac{\rho-1}{\rho+1}$	$\overline{\cos\phi}$	ϕ_b	$S_\rho/\overline{\cos\phi}$		$\cos\phi_b/\overline{\cos\phi}$.	
				Observed.	Calculated	Observed	Calculated.
2·11	0·357	0·193	75·7° ($\cos\phi_b$ = 0·247)	1·85	1·88	1·28	1·25

Asymmetry factor σ *calculated from*—

ρ.	$\overline{\cos\phi}$.	ϕ_b
1·39	1·38	1·34

According to the equations given in an earlier section the following relations exist between ρ, $\overline{\cos\phi}$ and ϕ_b.—

$$(\rho - 1)/(\rho + 1) = 15/8 \,.\, \overline{\cos\phi}$$
$$\cos\phi_b = 5/4 \,.\, \overline{\cos\phi}$$

Table VI shows that the experimental observations are in satisfactory agreement with the values required by the above equations. It thus appears that as far as the relative values of $\rho \,.\, \overline{\cos\phi}$ and ϕ_b are concerned, the scheme used in defining σ is justified by the experimental results. The final values of the asymmetry factor, σ, corresponding to the mean values of ρ, $\overline{\cos\phi}$ and ϕ_b are 1·39, 1·38 and 1·34 respectively, the mean being 1·37. The theoretical distribution curve corresponding to this value of σ is given in fig. 2, and its relation to the experimentally observed photoelectron distribution has already been discussed. The only differences which seem to be greater than statistical fluctuations are the more rapid decrease in the observed intensity on the smaller angle side of the maximum, and the slightly greater intensity of emission of photoelectrons at small angles with the incident rays

[*Note added October 5th.*—Since writing this paper we have seen an interesting note by Auger ('Comptes Rendus,' vol. 186, p 758, 1928), in which he gives the results of an examination of the initial directions of emission of the photoelectrons produced by tungsten K radiation, $\lambda = 0 \cdot 21$ Å Auger finds that the longitudinal asymmetry of the distribution corresponds to an average forward momentum of the photoelectrons about 30 per cent. greater than that of an absorbed quantum. This value is much greater than that previously obtained by Auger, but is in substantial agreement with the results obtained by the authors in the work described in the above paper.]

Summary.

A survey is given of previous observations on the longitudinal asymmetry and dispersion of the photoelectron distribution produced by X-rays, with reference to the requirements of modern theory. The possible sources of error which may explain the discordant results of previous observers are discussed and the precautions taken to eliminate them in the present investigation are described.

The results of the authors' own investigation given in this paper refer to the photoelectron emission from oxygen and nitrogen due to the absorption of X-rays of wave-lengths $0 \cdot 54$, $0 \cdot 61$ and $0 \cdot 71$ Å respectively. They show (1) that the spread or dispersion of the photoelectrons may be satisfactorily represented by the $\cos^2 \theta$ law, and (2) that the longitudinal asymmetry of the distribution corresponds to an *average* forward momentum of the photoelectrons nearly 40 per cent. greater than that of an absorbed quantum.

The authors wish to express their thanks to Prof. W. L. Bragg, F.R S, for his continued interest in the work and for many valuable suggestions during the course of its progress.

The Effect of the Image Force on the Emission and Reflexion of Electrons by Metals.

By Dr. L. W. NORDHEIM.

(Communicated by R. H. Fowler, F.R.S.—Received August 8, 1928.)

§ 1. *Introduction.*

In a recent paper* (cited henceforth as I) I have shown that the phenomena concerning the emission or reflexion of electrons by metals can be treated by calculating the emission or reflexion coefficient for the electrons at the surface of the metal and integrating over all incident electrons according to the electron theory of conductivity of Sommerfeld. In a further paper† (cited henceforth as II), R. H. Fowler and the present author have treated the cold emission in intense electric fields on the same principle.

The surface of the metal is characterised thereby as a region with a very sudden variation of the potential, that, according to the wave mechanics causes a reflexion. The emission coefficient is, of course, the ratio of the number of electrons going through to the number of incident electrons, and the relation

$$R + D = 1 \tag{1}$$

(R = reflexion coefficient, D = emission coefficient) is therefore always valid.

In the papers mentioned above, R and D have been calculated for the idealised forms of the potential steps A and B of fig. 1, C denoting the total height of the potential step.

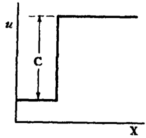


FIG. 1A.

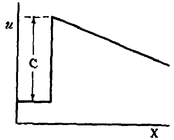


FIG. 1B.

---

* Nordheim, 'Z. Physik,' vol. 46, p. 833 (1928). Compare also Fowler, 'Roy. Soc. Proc.,' A, vol. 117, p. 549 (1928), and Sommerfeld, 'Z. Physik,' vol. 47, p. 1 (1928).

† Fowler and Nordheim, 'Roy. Soc. Proc.,' A, vol. 119, p. 173 (1928).

The real form of the potential is, of course, a smooth curve, and it is the purpose of this paper to give a more accurate calculation of D and R for a field that approaches much closer to the real circumstances.

The result for the intense field effect is a small decrease of the necessary field strengths, whereas the emission coefficient for the thermions, which was estimated in I at about $\frac{1}{2}$, becomes very nearly equal to 1, so that the average reflexion in that case amounts only to a few per cent.

§ 2. *The Field and the Differential Equation Employed.*

The chief deviation of the real fields from those of fig. 1 consists in the image field, which must give the right kind of rounding off of the upper corner of the potential curves of fig 1.* For our purposes, furthermore, the form of the upper part is of much greater importance than the lower part, since according to I (formula (32)) the expression, which measures the reflexion in a given interval, contains $(W - U)^{-3/2}$ as factor.

W denotes here the energy of the normal component of the velocity in the interior of the metal, the other components being of no importance, and U the potential energy. Therefore the regions, where W — U is small, give a much larger effect, whereas at great velocities of the electrons irregularities of the potential do not matter so much.

We have, therefore, quite a good approximation for our purpose by taking the image potential right down to the bottom of the total potential step, and connecting it there with the constant potential in the interior of the metal.

The above consideration justifies also the treating of the problem as the refraction of the de Broglie waves in a continuous medium, since the wave-length, $\lambda = h/mv$, is large compared with the atomic distances for the critical region with small W — U.

Now the image potential is equal to $- e^2/4x$, where x is the distance from the metal surface, and we take as the total potential therefore

$$U = C - e^2/4x - Fx \qquad \text{for } x > x_0 \tag{2A}$$

$$U = 0 \qquad \text{for } x < x_0, \tag{2B}$$

where x_0 is given by

$$e^2/4x_0 = C, \tag{3}$$

---

* An experimental proof of this is furnished by the good agreement of the Schottky correction (Schottky, 'Z. Physik,' vol. 14, p. 63 (1923)) for the influence of not too strong fields on the thermionic emission, which has been verified by various investigators. See, for instance, Pforte, 'Z. Physik,' vol. 49, p. 333 (1928), and de Bruyne, 'Roy. Soc. Proc.,' A, vol. 120, p. 423 (1928).

and we have thus the following graphs in the cases $F = V$ (fig. 2A) and $F \neq V$ (fig. 2B).

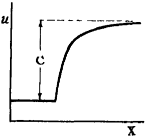


FIG 2A.

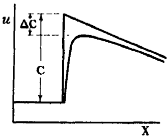


FIG 2B.

The value of C itself, *i.e.*, the total potential difference between the inside and the outside of the metal, can be estimated quite well as the sum of the partial potential

$$\mu = \frac{h^2}{8m}\left(\frac{6n}{\pi g}\right)^{2/3}, \tag{4}$$

where n is the number of free electrons per cubic centimetre, and $g = 2$ the statistical weight of quantum state of an electron, and the thermionic work function χ. This has been checked quite well by observing the refractive index for the de Broglie waves in the experiments about their Laue diffraction.*

Now for one ionised electron per atom μ amounts to already about 8 volts and for two electrons per atom to 12 volts, and $C = \mu + \chi$ may vary therefore between 10 and 20 volts for various metals.

We consider at first only the case $F = 0$, *i.e.*, without an external field. We have then to solve the wave equations

$$\frac{d^2\psi}{dx^2} + \kappa^2\left(W - C + \frac{e^2}{4x}\right)\psi = 0 \quad \text{for } x > x_0, \tag{4A}$$

$$\frac{d^2\psi}{dx^2} + \kappa^2 W\psi = 0 \qquad\qquad \text{for } x < x_0, \tag{4B}$$

$$\kappa^2 = 8\pi^2 m/h^2$$

under the condition that both, ψ and $d\psi/dx$ are continuous at $x = x_0$ and ψ represents asymptotically for $x \gg x_0$ a stream of outgoing electrons.

The problem is therefore to find that special solution of (4A) for large x and to obtain a development of it in the neighbourhood of x_0, which is small.

---

* See, for instance, the summary given by Rosenfeld and Witmer, 'Z. Physik,' vol. 49, p. 534 (1923).

This can be done by the integral of Barnes,* which allows us to connect the solutions for large and small x.

§ 3. *The Solution of the Differential Equation.*

By the substitution

$$x = \tfrac{1}{2}z\,(\mathrm{W}-\mathrm{C})^{-\frac{1}{2}}/\kappa, \tag{5}$$

the differential equation (4A) becomes

$$\frac{d^2\psi}{dz^2} + (\tfrac{1}{4} + \tfrac{1}{z}k)\psi = 0, \tag{6}$$

where

$$k = \tfrac{1}{8}\kappa e^2\,(\mathrm{W}-\mathrm{C})^{-\frac{1}{2}},\ \ z_0 = e^2\kappa\,(\mathrm{W}-\mathrm{C})^{\frac{1}{2}}/2\mathrm{C}. \tag{7}$$

The magnitude in powers of which the final development will proceed is

$$s = z_0 k = \kappa^2 e^4/16\mathrm{C}, \tag{8}$$

which is for $\mathrm{C} = 2 \,.\, 10^{-11}$ (approximately 12 volts) equal to 0·25.
Now putting

$$\psi = e^{-\frac{1}{2}iz} z^{-ik}\, v, \tag{9}$$

we get as equation for v

$$z^2\frac{d^2v}{dz^2} - 2ikz\frac{dv}{dz} + ik\,(ik+1)\,v - iz^2\frac{dv}{dz} = 0,$$

or changing the variable

$$z = -it, \tag{10}$$

$$t^2\frac{d^2v}{dt^2} - 2ikt\frac{dv}{dt} + ik\,(ik+1)\,v - t^2\frac{dv}{dt} = 0 \tag{11}$$

Of course a second solution of (6) is obtained by taking a positive sign in the exponential function in (9), but the above solution represents the outgoing waves that we need.

Equation (11) is now satisfied by the integral

$$\mathrm{I} = \int_{-\infty i}^{+\infty i} \Gamma(s)\Gamma(-s+ik)\Gamma(-s+ik+1)\,t^s\,ds. \tag{12}$$

---

* For a full development of this method see Whittaker and Watson, 'Modern Analysis,' Cambridge, 4 ed., 1927, p. 243; also all other references in respect of analytical matters will be given to this treatise.

The poles of the integrand and the path of integration are indicated in fig 3, and the validity of (12) is shown as follows

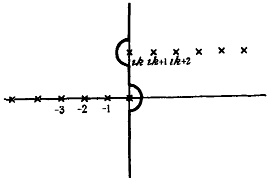


Fig. 3.

Substituting I for v in (11) we get for the left side

$$\int_{-\infty i}^{+\infty i} \Gamma(s)\Gamma(-s+ik)\Gamma(-s+ik+1)\left[s(s-1)t^{s}-2ikst^{s}+ik(ik+1)t^{s}-st^{s-1}\right]ds,$$

which is equal to

$$\int_{-\infty i}^{+\infty i}\{\Gamma(s)\Gamma(-s+ik+1)\Gamma(-s+ik+2)\,t^{s}$$
$$-\Gamma(s+1)\Gamma(-s+ik)\Gamma(-s+ik+1)\,t^{s+1}\}\,ds$$
$$-\left(\int_{-\infty i}^{+\infty i}-\int_{1-\infty i}^{1+\infty i}\right)\Gamma(s)\Gamma(-s+ik+1)\Gamma(-s+ik+2)\,t^{s}\,ds.$$

Since there are no poles of the last integrand between the paths of integration, and the integrand tends to zero as $|s|\to\infty$, the last expression vanishes according to Cauchy's theorem, and (12) is, therefore, a solution of (11). The behaviour of the above integrands for large (s) can thereby be estimated by the asymptotic expression of the gamma function (*cf.* Whittaker and Watson, p 278).

$$\Gamma(s+a)\to s^{s+a-\frac{1}{2}}\,e^{-s}O(1). \tag{13}$$

The integrand in (12) is accordingly for large $|s|$ of the order of magnitude

$$|s|^{-s+2ik-\frac{1}{2}}\,e^{s}t^{s}$$

It tends to zero for all large s with positive real part, but finite t, and one can therefore transform the contour of integration into small circles around the poles on the right side, and it is thus

$$\mathrm{I}=-2\pi i\,\Sigma \text{ Residues at } s=ik,\ \ ik+1,\ \ ik+2,\ \ldots.$$

This gives a series for v that converges for all finite values of t.

On the other hand for every s on the left side of the i-axis a t can be found so that the integrand in (12) becomes as small as desired, and this can be used to obtain an asymptotic expansion for large t by using a transformation of the contour of integration as indicated in fig. 4. This is, of course the analytical continuation of the solution for finite t, and allows one thus to study the behaviour of the latter near infinity

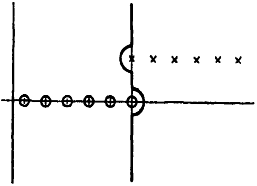


Fig. 4

The asymptotic solution has the form

$$I = 2\pi i \, \Sigma \text{ Residues at } s = 0, -1, -2, \quad r$$

plus a remainder R_r, which can be made as small as desired for large enough k's and as

$$\operatorname*{Res}_{z=-n} \Gamma(z) = \frac{(-1)^n}{n!}, \quad (n = \text{positive integer})$$

we obtain

$$v = I = 2\pi i \sum_0^r \frac{(-1)^n}{n!} t^{-n} \, \Gamma(n + ik) \, \Gamma(n + ik + 1) + R_r$$

$$= 2\pi i \Gamma(ik) \, \Gamma(ik + 1) \sum_0^r (-1)^n t^{-n} [ik(ik + 1)] \ldots [(ik + n - 1)(ik + n)]/n! + R_r,$$

i.e., a single expansion in descending powers of t, which shows that (9) really represents outgoing waves.

The expansion for finite values of t becomes, always omitting irrelevant constant factors,

$$v = \Sigma \operatorname*{Res}_{s = ik, ik+1, ik+2,} F \tag{15}$$

where

$$F = \Gamma(s) \, \Gamma(-s + ik) \, \Gamma(-s + ik + 1) \, t^s. \tag{16}$$

The pole $s = ik$ is a simple one and has the residue

$$\operatorname*{Res}_{s = ik} \mathrm{F} = - \Gamma(ik)\, t^{ik}. \tag{17}$$

The poles $s = ik + 1, ik + 2, \ldots$ are of the order two.

With the help of the relation (*cf.* Whittaker and Watson, p. 239)

$$\Gamma(z)\, \Gamma(1 - z) = \pi / \sin \pi z,$$

F can be written

$$\mathrm{F} = - \pi^2 g / \sin^2 \pi (s - ik)\,; \quad g = \Gamma(s)\, t^s / \Gamma(1 + s - ik)\, \Gamma(s - ik) \tag{18}$$

where g is regular for the values of s considered. Now we have the development in rational fractions

$$\pi^2 / \sin^2 \pi z = \sum_{-\infty}^{+\infty} (z - n)^{-2},$$

and the residue of F becomes therefore

$$\operatorname*{Res}_{s = ik + n} \mathrm{F} = [- dg/ds]_{s = ik + n}. \tag{19}$$

From (15), (17), (18) and (19) we obtain (omitting the factor $-\Gamma(ik)$)

$$\begin{aligned} v &= t^{ik} + \sum_{n=1}^{\infty} \left[\frac{d}{ds} \left(\Gamma(s)\, t^s / \Gamma(ik)\, \Gamma(1 + s - ik)\, \Gamma(s - ik) \right) \right]_{s = ik + n} \\ &= t^{ik} \left(1 + \sum_{n=1}^{\infty} \left[\frac{d}{dz} \left(\Gamma(z + ik)\, t^z / \Gamma(ik)\, \Gamma(1 + z)\, \Gamma(z) \right) \right]_{z = n} \right) \end{aligned} \tag{20}$$

By repeated application of the formula

$$\frac{\Gamma'(z + a)}{\Gamma(z + a)} = \frac{1}{z + a - 1} + \frac{\Gamma'(z + a - 1)}{\Gamma(z + a - 1)},$$

which is obtained by logarithmic differentiation of

$$\Gamma(z + a) = (z + a - 1)\, \Gamma(z + a - 1),$$

one gets

$$\left[\frac{d}{dz} \left(\frac{1}{\Gamma(z)} \right) \right]_{z = n} = - \frac{1}{\Gamma(n)} \left(- \frac{1}{n} + \frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{n} - \gamma \right), \quad n = 1, 2, \ldots \tag{21A}$$

$$\left[\frac{d}{dz} \left(\frac{1}{\Gamma(z + 1)} \right) \right]_{z = n} = - \frac{1}{\Gamma(n + 1)} \left(1 + \frac{1}{2} + \ldots + \frac{1}{n} - \gamma \right), \tag{21B}$$

$$[\Gamma'(z + ik)]_{z = n} = \Gamma(n + ik) \left(\frac{1}{ik} + \frac{1}{ik + 1} + \ldots + \frac{1}{ik + n - 1} + \frac{\Gamma'(ik)}{\Gamma(ik)} \right), \tag{21C}$$

$\gamma = \Gamma'(z)/\Gamma(z)$ for $z = 1$ being Euler's constant ($= 0{\cdot}5772 \ldots$).

Carrying out the differentiation in (20) with the help of these formulæ one obtains

$$v = t^{ik}\left[1 + \sum_{n=1}^{\infty} t^n \mathrm{A}_n (\log t + \mathrm{B}_n^*)\right], \tag{22}$$

where

$$\mathrm{A}_n = \Gamma(n+ik)/\Gamma(ik)\,\Gamma(n+1)\,\Gamma(n) = ik(ik+1)\,.\,.\,(ik+n-1)/n!\,(n+1)! \tag{23A}$$

$$\mathrm{B}_n^* = 2\gamma + \frac{1}{n} - 2\left(\frac{1}{1} + \frac{1}{2} + \ldots \frac{1}{n}\right) + \frac{1}{ik} + \frac{1}{ik+1} + \ldots \frac{1}{ik+n-1} + \frac{\Gamma'(ik)}{\Gamma(ik)} \tag{23B}$$

Returning to the original variable $z = -it$ (*cf.* (10)), we obtain from (9) (omitting again i^{-ik} in the second row) as the final result

$$\begin{aligned}\psi &= e^{-\frac{1}{2}iz} z^{-ik} v(iz) \\ &= e^{-\frac{1}{2}iz}\left[1 + \sum_{n=1}^{\infty} (iz)^n \mathrm{A}_n (\log z + \mathrm{B}_n)\right],\end{aligned} \tag{24}$$

where

$$\mathrm{B}_n = \mathrm{B}_n^* + \log i = \mathrm{B}_n^* + \tfrac{1}{2}i\pi. \tag{25}$$

It may be remarked that

$$\psi = e^{-\frac{1}{2}iz} \sum_{n=1}^{\infty} (iz)^n \mathrm{A}_n \tag{26}$$

is a second independent solution of (6), which may be obtained by the ordinary series method, but, of course, does not represent outgoing waves. The method used above offers, incidentally, also quite a convenient way of discussing the solutions of the Schrödinger equation for the hydrogen atom.

To complete our solution we have only to obtain a suitable expression for $\Gamma'(ik)/\Gamma(ik)$. Since, according to (7), k is becoming very large in the neighbourhood of $\mathrm{W} = \mathrm{C}$, we need an expansion for large k. As it is not allowed to differentiate asymptotic expansions directly, we have to proceed in the following way. From the general formula

$$\Gamma'(z)/\Gamma(z) = -\gamma - \tfrac{1}{z} + z\sum_{1}^{\infty} n^{-1}(z+n)^{-1},$$

(Whittaker and Watson, p. 241) we obtain

$$\Gamma'(ik)/\Gamma(ik) = -\gamma + i/k + \psi + i\phi, \tag{27}$$

where

$$\psi = k^2 \sum_{1}^{\infty} n^{-1}(k^2+n^2)^{-1}; \quad \phi = k\sum_{1}^{\infty} (k^2+n^2)^{-1}. \tag{28}$$

These series can be summed by Euler's formula (compare Whittaker and Watson, p. 128)

$$f(n)+f(n+1)+\ldots f(m)=\int_n^m f(x)\,dx+\frac{f(m)+f(n)}{2}$$
$$+\sum_{r=1}^{s}(-1)^{r-1}\frac{B_r}{(2r)!}\left(f_{(m)}^{(2r-1)}-f_{(n)}^{(2r-1)}\right)+R_s \qquad (29)$$

n and m are here positive integers and $m > n$; B_r are the Bernoullian numbers ($B_1 = 1/6$), and the remainder is of the order of magnitude

$$R_s = V\left(\int_n^m f^{(2s+1)}(x)\,dx\right).$$

Taking $n = 1$, $m = \infty$, $f(x) = k^2/x(k^2 + x^2)$ we obtain

$$\psi=\int_1^\infty\frac{dx}{x(1+x^2/k^2)}+\frac{k^2}{2(k^2+1)}+\frac{1}{12}\left[\left(-\frac{1+x^2/k^2+2x^2/k^2}{x^2(1+x^2/k^2)^2}\right)\right]_1^\infty+\ldots$$
$$=\left|\log z-\tfrac{1}{2}\log(1+z^2)\right|_{1/k}^{\infty}+\tfrac{1}{2}-\frac{1}{2k^2}+\frac{k^2+3}{12(k^2+2+1/k^2)}+\ldots$$
$$=\log k+\tfrac{7}{12}+\frac{1}{12(k^2+2)}+\ldots. \qquad (30)$$

Similarly ϕ becomes

$$\phi=\int_1^\infty\frac{dx}{k(1+x^2/k^2)}+\frac{1}{2(k+1/k)}+\ldots$$
$$=\Big[\arctan z\Big]_{1/k}^{\infty}+\frac{1}{2k}+\ldots=\tfrac{1}{2}\pi-\frac{1}{2k}+0\left(\frac{1}{k^3}\right), \qquad (31)$$

and therefore the required asymptotic expression is

$$\frac{\Gamma'(ik)}{\Gamma(ik)}=\log k-\gamma+\tfrac{7}{12}+i\left(\tfrac{1}{2}\pi+\frac{1}{2k}\right)+\ldots, \qquad (32)$$

which gives according to (23B) and (25)

$$B_n=+\gamma+\frac{7}{12}+\frac{1}{n}-2\left(\frac{1}{1}+\frac{1}{2}+\ldots\frac{1}{n}\right)+\log k+\frac{1}{ik}$$
$$+\frac{1}{ik+1}+\ldots\frac{1}{ik+n-1}+i\left(\pi+\frac{1}{2k}\right)+\ldots. \qquad (33)$$

§ 4 *Calculation of the Emission or Reflexion Coefficient.*

Taking as solution for (4B)

$$\psi_a^{*}=ae^{-ik(x-x_0)\sqrt{W}}+a'e^{+ik(x-x_0)\sqrt{W}}, \qquad (34)$$

the continuity conditions for ψ and $d\psi/dx$ at x_0 are

$$a + a' = \psi(z_0)$$

$$ik(-a + a')\,W^{\frac{1}{2}} = \left[\frac{d\psi}{dz}\frac{dz}{dx}\right]_{x_0} = 2k\,(W - C)^{\frac{1}{2}}\left[\frac{d\psi}{dz}\right]_{z_0}$$

Now, since according to (24)

$$\frac{d\psi}{dz} = -\frac{i}{z}e^{-iz/2}\left[1 + \sum_{n=1}^{\infty}(iz)^n A_n(\log z + B_n) - 2\sum_{n=1}^{\infty}(iz)^{n-1}A_n(1 + n\log z + nB_n)\right]. \tag{35}$$

the above relations can be written in the form

$$\left.\begin{aligned} a + a' &= S \\ -a + a' &= -\rho T \end{aligned}\right\}, \tag{36}$$

where

$$\left.\begin{aligned} S &= 1 + \sum_{1}^{\infty}(iz)^n A_n(\log z + B_n) \\ T &= S - 2\sum_{1}^{\infty}(iz)^{n-1}A_n(1 + n\log z + nB_n) \end{aligned}\right\}, \tag{37}$$

$$\rho = \{(W - C)/W\}^{\frac{1}{2}}, \tag{38}$$

and the emission coefficient becomes

$$D(W) = 1 - \frac{|a'|^2}{|a|^2} = \frac{2\rho(S\bar{T} + \bar{S}T)}{|S|^2 + \rho^2|T|^2 + \rho(S\bar{T} + \bar{S}T)}. \tag{39}$$

The evaluation of this expression with the help of (23A), (33) and (37) is rather tiresome but quite elementary. All terms can be expressed as functions of

$$X = \{(W - C)/C\} \quad \text{and} \quad s = z_0 k = k^2e^4/16C, \tag{40}$$

the second of which is for $C = 6 \,.\, 10^{-12}$ equal to $0{\cdot}25$, and is the only constant that depends on the special metal. For discussing the results it is convenient to expand numerator and denominator in (39) in powers of X, the coefficients of which are itself series in S. Neglecting higher powers than X^2 and S^2 we obtain as the final result

$$D(W) = (a + bX + cX^2 + \ldots)/(d + eX + fX^2 + \ldots) = \alpha + \beta X + \gamma X^2 + \ldots, \tag{41}$$

where

$$\left.\begin{aligned} &a = 4\pi\sqrt{s}\,;\ b = 4\,.\,16{\cdot}49\,.\,s^2,\ c = -2\pi\sqrt{s}\,(1 + 4s) \\ &d = 1 + s\,[(\log s)^2 + 0{\cdot}311\log s + 10{\cdot}90] \\ &\qquad\qquad + s^2\,[-(\log s)^2 + 0{\cdot}689\log s - 9{\cdot}23] \\ &e = -6{\cdot}28\sqrt{s}\,(1 + s) + \tfrac{1}{2}(a + b) \\ &f = 2 - s\,[(\log s)^2 + 8{\cdot}321\log s + 14{\cdot}18] \\ &\qquad\qquad + s^2\,[-(\log s)^2 + 1{\cdot}96\log s + 9{\cdot}23] + \tfrac{1}{2}(a + b) \end{aligned}\right\}, \tag{42}$$

and

$$\alpha = a/d\,;\ \beta = (bd - ea)/d^2\,;\ \gamma = (cd^2 - efd - afd + e^2a)/d^3. \tag{43}$$

The influence on the thermionic emission is given by the average emission coefficient (compare I (17)).

$$\bar{\mathrm{D}} = \frac{1}{k\mathrm{T}}\int_{\mathrm{C}}^{\infty} \mathrm{D}\,(\mathrm{W})\, e^{-(\mathrm{W}-\mathrm{C})/k\mathrm{T}}\, d\mathrm{W} = \int_{0}^{\infty} \mathrm{D}\,(k\mathrm{T}z)\, e^{-z}\, dz$$

$$= \int_{0}^{\infty} (\alpha + \beta\sqrt{k\mathrm{T}/\mathrm{C}}\sqrt{z} + \gamma\, k\mathrm{T}/\mathrm{C}\, z + \ \ldots) e^{-z}\, dz$$

$$= \alpha + 0{\cdot}886\,\beta\sqrt{k\mathrm{T}/\mathrm{C}} + \gamma\, k\mathrm{T}/\mathrm{C} + \ldots \qquad (44)$$

The expansions (41) or (44) are, of course, valid only for sufficiently small $\mathrm{X} = (\mathrm{W}-\mathrm{C})/\mathrm{C}$ or $k\mathrm{T}/\mathrm{C}$, where $\mathrm{W}-\mathrm{C}$ is the energy of the escaping electrons and C the total amount of the potential difference on the surface.

For the special value $\mathrm{C} = 2 \,.\, 10^{-11}$, *i.e.*, $s = 0{\cdot}24$, we get the following numerical values,

$$\alpha = 0{\cdot}927 \qquad \beta = 0{\cdot}431 \qquad \gamma = -1{\cdot}018.$$

The emission coefficient does not tend therefore to zero for vanishing escaping velocity, but only to the already very high value 0·927 At the same time the reflexion coefficient $\mathrm{R} = 1 - \mathrm{D}$ for electrons, hitting the surface from the outside, is only about 0·07 for the slowest electrons, and it seems, therefore, hardly possible to observe this effect.

Furthermore it follows from this calculation that the coefficient A in the Richardson-Dushmann formula for thermionic emission

$$i = \mathrm{A}\mathrm{T}^{2} e^{-b/\mathrm{T}}, \qquad (45)$$

which is according to the theory (compare I (18))

$$\mathrm{A} = 2\pi m e k^{2} g \bar{\mathrm{D}}/h^{3} \qquad (46)$$

has a value near 120 and not 60 amp /cm.², which is obtained by neglecting the factor $g\bar{\mathrm{D}}$ ($g = 2$ being the weight factor for the electronic quantum states due to the spin). But it seems exceedingly difficult, owing to the overwhelming influence of the exponential function, to determine A with sufficient accuracy to decide between those values experimentally.

The difference between the results of this more accurate determination of D and R and my former estimation of it is, of course, due to the very slow vanishing of the image field for large distances, which causes the emission coefficient to approach a non-zero limit instead of a limit zero as at a sharp potential step.

§ 5. *The Case of Intense Electric Fields.*

We consider now the case of an external field. The rigorous treatment of it would demand a complete discussion of the differential equation (compare 4A))

$$\frac{d^2\psi}{dx^2} + k^2\left(W - C + \frac{e^2}{4x} + Fx\right)\psi = 0, \tag{47}$$

which has not yet been done in the mathematical literature, as far as I know. (47) is, of course, a special case of the equation for the Stark effect, as (4A) is for the hydrogen atom, but existing investigations do not give what we want here, and it seems to be very difficult to obtain a complete solution. But for values of W that are rather smaller than C, a method of taking account of the image field has already been suggested in II, § 4. The most important part of the equation for emission in strong electric fields (II (18) and (21)) is the exponential factor, which we determined, using the field of fig. 1 (ii) to be

$$e^{-2Q} = \exp\left[-4\kappa\,(C - W)^{3/2}/3F\right]. \tag{48}$$

Now, applying the method of Jeffreys (compare I, § 4), one finds that for an arbitrary potential of a shape like that of fig 5, the exponent becomes

$$2Q = 2\kappa\int_{x_1}^{x_2}(U - W)^{\frac{1}{2}}\,dx, \tag{49}$$

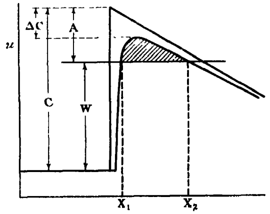


FIG. 5.

the integration being taken over the shaded area in fig. 5. For our potential (2A) we have therefore

$$2Q = 2\kappa\int_{x_1}^{x_2}(C - W - e^2/4x - Fx)^{\frac{1}{2}}\,dx, \tag{50}$$

x_1 and x_2 being the zero points of the integrand, and, therefore, putting $C - W = A$

$$x_{2,1} = \{A \pm \sqrt{(A^2 - e^2F)}\}/2F. \tag{51}$$

The integral in (50) is a simple elliptic integral that can be transformed in the usual way to the following normal form

$$2Q = \frac{2\sqrt{2}}{3F} A \sqrt{\{A^2 + \sqrt{(A^2 - e^2F)}\}} \left[E(k) - \frac{F(e^2)}{A\{A + \sqrt{(A^2 - e^2F)}\}} K(k) \right], \tag{52}$$

where E and K are the elliptic normal integrals*

$$K(k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \phi)^{-\frac{1}{2}} \, d\phi, \quad E(k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \phi) \, d\phi,$$

and

$$k = 2\sqrt{(A^2 - e^2F)}/\{A + \sqrt{(A^2 - e^2F)}\}. \tag{53}$$

The emission formula will become to a sufficient approximation (compare II (21))

$$I = \text{const } F^2 e^{-2Q}, \tag{54}$$

where in the formula (52) for Q one has to take the thermionic work function χ as the value for A. According to this formula, of course, log I will not give exactly a straight line when plotted against $1/F$, but the deviation is so small that it will most probably not be possible to observe it.

For the discussion and comparison of (54) with our former calculation it may be noted, that $e\sqrt{F} = \Delta C$ is the drop in the work function effected by the external field, as pointed out first by Schottky. The quotient V of the new value of Q (52) and the former one (compare (43)) is

$$V = \sqrt{\left(\frac{1 + \sqrt{(1 - x^2)}}{2}\right)} \left(E(k) - \frac{x^2}{1 + \sqrt{(1 - x^2)}} K(k)\right). \tag{55}$$

$$x = e\sqrt{F}/\chi; \quad k = 2\sqrt{1 - x^2}/(1 + \sqrt{1 - x^2}),$$

V that measures directly the influence of the image force on the cold electronic emission, depends only on x, *i.e.*, the ratio of the Schottky correction to the undisturbed work function, and the values of it are given by the following table :—

X	0	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1
V	1	0·951	0·904	0·849	0·781	0·696	0·603	0·494	0·345	0

---

* Tables of the elliptic integrals are given, for instance, in 'Jahnke-Emde, Funktionentafeln,' Leipzig, 1909, p. 68.

and the emission formula can be written with sufficient approximation

$$I = \text{const}\; e^{-4k\chi^{3/2}v/F}. \tag{56}$$

If we measure χ in volts and F in volts per centimetre we obtain

$$x = 3\cdot 78 \,.\, 10^{-4}\sqrt{F}/\chi.$$

For $\chi = 4\cdot 5$ volts, which corresponds to pure tungsten, and $F = 3 \,.\, 10^7$, at which figure the emission becomes sensible (according to our formula, II (22)), x becomes $0\cdot 46$, and the v therefore about $0\cdot 8$, which means a corresponding reduction for the necessary field strengths. That reduction, of course, is far from sufficient to bring the coefficient of $1/F$ down to the observed values and one has still to assume surface irregularities or sensitive spots.

The transition from the cold emission to the thermionic emission could be treated rigorously only by solving equation (47) for values of W that are a little smaller but very nearly equal to C. But it appears quite clearly from the present calculations that the two effects will be fairly well additive. Since nearly all electrons with a larger energy than the potential threshold can escape, it is to be expected that the Schottky correction for thermionic emission holds quite well in conformity with the experimental results. On the other hand the cold emission is caused by the much larger number of slower electrons according to the Sommerfeld distribution function of the electrons in the metal that can escape only under the influence of the external field.

I am greatly indebted to Dr. R. H. Fowler for his interest and advice and to the International Education Board for enabling me to work at Cambridge.

The Thermal and the Electrical Conductivity of a Copper Crystal at Various Temperatures.

By W. G. Kannuluik, B Sc , and Prof. T. H. Laby, M.A., Sc.D., F.Inst.P.,
Natural Philosophy Laboratory, University of Melbourne

(Communicated by Sir Ernest Rutherford, P R S.—Received August 8, 1928.)

Summary.

The thermal conductivity of a number of regular and non-regular single metallic crystals has been studied by previous workers. There appear to be no marked differences between the conductivity of single and polycrystal specimens of the regular metals, except at low temperatures, but further confirmation of this statement by additional accurate investigations seems necessary, more particularly at low temperatures, where different observers do not agree.

The methods which have been used to determine the thermal conductivity of metals are briefly discussed ; the thermal method appears to be the simplest, but the electrical or Kohlrausch's method is better adapted for very low temperatures, *i.e.*, near 20° K.

The authors have used the thermal method for determining the thermal conductivity of a copper crystal over a range of temperatures. The heat losses have been made small by using a high vacuum, and then at the room temperature eliminated by a method the theory of which is given below.

The thermal conductivity, λ, is found to be 0·989 cal. cm. sec. deg. at 19·4° C., 1·054 at $-$ 73·7° C , and 1·131 at $-$ 174·8° C. The value of λ at 19·4° C. is about 4 per cent. greater than that of polycrystal copper (in agreement with Schott's work), while the values of λ obtained at the lower temperatures are considerably less than Schott's values. The electrical conductivity, κ, of the single crystal proved to be the same as that of ordinary polycrystal copper. The values of the Wiedemann-Franz constant $\lambda/(\kappa T)$ obtained by us are 1 69 $\times$ 10^{-8} watt ohm deg.$^{-2}$ at 90·2° K., and 2·42 $\times$ 10^{-8} at 273·2° K.

Introduction.

Previous Work on Single Crystals of the Metals.—The determination of the thermal and the electrical properties of the metals is desirable on general grounds. The single crystal probably represents the thermodynamically most stable form of a metal, provided it is pure and free from mechanical strain.

To what extent the thermal and the electrical properties of the usual polycrystal form are dependent on the crystallite structure has not yet been cleared up. According to Eucken and Neumann,* the thermal conductivity of the poor metal conductors, antimony and bismuth, is considerably greater in coarse-grained specimens than in fine, while the corresponding variations in the electrical conductivity were, in comparison, quite small. On the other hand, the recent researches of Grüneisen and Goens† on a large number of specimens (polycrystal and single crystal) of the regular metals, copper, gold, aluminium and platinum, at —190° C. and —252° C. seem to prove that impurities and, to a lesser extent, cold work (*e.g.*, hammering) diminish both conductivities. They leave it an open question whether there is or is not any extra resistance to the flow of heat or electricity at the common boundaries of adjacent crystallites.

The actual amount of single crystal data of the thermal conductivity is not large. Besides the considerable contributions to it by Gruneisen and Goens, there have been a number of other more restricted investigations. Schott‡ has obtained the conductivity of a natural copper crystal from --252° C. to 0° C.; Kaye and Roberts§ that of bismuth crystals at room temperature; Bridgman‖ the conductivity of a number of ånisotropic metal crystals at 20° C.; and Ezer Griffiths¶ that of an aluminium crystal at room temperature. Griffiths concluded that the thermal conductivity of a single crystal rod of aluminium was no greater than that of an extruded polycrystal rod of the same purity. A single determination of the single crystal conductivity of the regular metals—those crystallizing in the cubic system—suffices, but it is necessary to make at least two determinations in different directions making known angles with the principal axis in the non-regular metal crystals. Bridgman and Kaye and Roberts give the conductivity in directions parallel and perpendicular to the principal axis of the crystal.

The values which have been found for the thermal conductivity of the metals, even from the more careful investigations, do not agree. The lack of agreement is least in the room temperature values, and becomes more marked the further removed the temperature is from room temperature. The discrepancies in the comparable values is due partly to experimental error and partly to differences in the purity and in the crystallite structure of the metals themselves. It is impossible to decide to what extent each of these factors contribute to the discrepancies. M. Jakob* has published a critical review of the polycrystal data of a number of the commoner metals and alloys. He has computed a probable error of 5 per cent. in the weighted mean value of λ for copper, silver, zinc and tin over the range −160° C. to 100° C. The probable error is certainly much greater outside this range. At very low temperatures, as at − 252° C., slight differences in the purity of different samples of the same metal (*e.g.*, copper) are responsible for very great variations in the conductivity. The large discrepancies in data at high temperatures must be put down to experimental shortcomings. It is especially difficult to allow for the heat losses, and to obtain the necessary precision with the temperature measurements with thermocouples at high temperatures (100° C. to 900° C.).

---

* 'Z. Phys. Chem.,' vol. 111, p. 473 (1924).

† 'Z. Phys.,' vol. 44, p. 615 (1927), and vol. 46, p 151 (1927).

‡ 'Ber. D. Phys. Ges,' vol. 18, p. 27 (1916).

§ 'Roy. Soc. Proc.,' A, vol. 104, p. 98 (1923).

‖ 'Proc. Amer. Acad.,' vol. 61, p. 101 (1926).

¶ 'Roy. Soc. Proc.,' A, vol. 115, p. 236 (1927).

Experimental Methods.—Careful investigations of the thermal conductivity of the metals have been made by Jaeger and Diesselhorst, Lees, W. Meissner, Schott, Grüneisen and Goens.† Of these, Jaeger and Diesselhorst applied Kohlrausch's well-known method, while Meissner used a modification of this method suitable for wires (1 mm. in diameter) over the range - 252° C. to −100° C He reduced the lateral heat losses to a minimum by using a high vacuum, and thereby achieved a better solution of the problem of heat-losses than did Jaeger and Diesselhorst. The remaining investigations quoted above are by the thermal method. One end of the metal rod is heated electrically and the temperature gradient is measured by thermocouples (in Lees' experiments by platinum thermometers). The correction for the lateral loss of heat is the chief difficulty of the thermal method. Schott, and after him Bridgman, and Grüneissen and Goens made the lateral loss of heat very small by using a high vacuum By ignoring the lateral heat loss altogether, the conductivities obtained by them were possibly in error by a few per cent. Except at very low temperatures, *circa* −250° C., the thermal method is satisfactory. At such temperatures the difference of the thermal E.M.Fs. of the thermocouples which determines the temperature gradient is so small in good conductors, *e.g.*, copper and gold, that an accurate determination of λ as in Gruneissen and Goen's experiments can scarcely be expected. It is here that the advantage of Meissner's method is apparent. The determination of the ratio λ/κ (κ the electrical conductivity) by this method is based ultimately on the measurement of the temperature coefficient of the resistance $\frac{1}{R}\frac{dR}{dT}$. This quantity *increases* as the temperature falls away, while the temperature gradient in the thermal method *diminishes*. On the other hand, the theory of the Schott method is the simplest possible, whereas in the region of the lowest temperatures that of Meissner is complicated to an altogether undesirable degree.

---

* 'Z Metallkde.,' vol 18, p. 55 (1926).

† Jaeger and Diesselhorst, 'Wiss. Abh. Phys. Tech. Reichsanst.,' vol. 3, p. 269 (1900); Lees, 'Phil. Trans.,' A, vol 209, p. 381 (1908); Meissner, 'Ann. Physik,' vol. 47, p. 1001 (1915); Schott, Grüneisen and Goens, *loc. cit.*

Theory.

The following theory is abbreviated from an unpublished investigation by one of the authors (W. G. K.) of the losses in the thermal method when λ is a linear function of the temperature.

In our experiments we have used a method apparently first used by Schott, but have developed the theory of the method sufficiently to allow for the very small lateral loss of heat.

The fundamental differential equation for the flow of heat along a rod is

$$\lambda q \frac{d^2\theta}{dx^2} = ph\,(\theta - \theta_s),^* \qquad (1)$$

where θ degree is the temperature at any section of the rod distant x from any convenient origin on its axis:

θ_s degree is the temperature of the surrounding medium;

q cm.^{2} the cross-section; p cm. the perimeter of the rod;

λ cal. cm. sec. deg. the thermal conductivity;

h cal. cm. sec. deg. the (Fourier) external conductivity.

The quantity h is of an exceedingly complex nature, involving as it does both the lateral loss from the rod by radiation and by conduction. The loss of heat from heated wires and rods in highly rarefied gases has been carefully investigated both theoretically and practically by Knudsen,† and also theoretically by Smoluchowski.‡ The justification of combining both losses into the single quantity h is in making the radiation loss small by silvering the radiating surfaces, and by so carrying out the experiments that the mean temperature of the rod is only a few degrees above that of its surroundings.

---

* Preston, 'Theory of Heat,' p. 607.

† 'Ann. Physik,' vol. 34, p. 593 (1911).

‡ 'Ann. Physik,' vol. 35, p. 983 (1911).

The equation (1) may be conveniently written in the form:

$$\frac{d^2z}{dx^2} = \mu^2 z, \tag{2}$$

where

$$\mu^2 \equiv ph/q\lambda \quad \text{and} \quad z \equiv \theta - \theta_s.$$

The solution of (2) is

$$Z = \mathrm{A} \sinh \mu x + \mathrm{B} \cosh \mu x \tag{3}$$

with the conditions that when

$$\begin{cases} x = 0 & \theta = \theta_1 \quad i.e., z = z_1 \\ x = l & \theta = \theta_0 \quad i.e., z = z_0. \end{cases}$$

On evaluating the integration constants A and B, and substituting their values in (3), we get after a slight reduction

$$z = \{z_1 \sinh \mu (l - x) + z_0 \sinh \mu x\}/\sinh \mu l,$$

which gives the distribution of temperature along the rod. The temperature gradient at any point is

$$\frac{d\theta}{dx} = \frac{dz}{dx} = \{-\mu z_1 \cosh \mu (l - x) + \mu z_0 \cosh \mu x\}/\sinh \mu l.$$

The gradients at the sections $x = 0$ and $x = l$ of the rod are respectively

$$\left.\begin{aligned} \left(\frac{d\theta}{dx}\right)_{x=0} &= \mu \{\theta_0 - \theta_s) \operatorname{cosech} \mu l - (\theta_1 - \theta_s) \coth \mu l\} \\ \left(\frac{d\theta}{dx}\right)_{x=l} &= \mu \{(\theta_0 - \theta_s) \coth \mu l - (\theta_1 - \theta_s) \operatorname{cosech} \mu l\} \end{aligned}\right\}, \tag{4}$$

In the evacuation method h can be made so small, and, therefore, μ also, that both coth μl and cosech μl may be replaced without sensible error by the first two terms in their expansions as power series. Using the well-known series

$$x \coth x = 1 + \tfrac{1}{3}x^2 - \tfrac{1}{45}x^4 + \ldots, \quad x \operatorname{cosech} x = 1 - \tfrac{1}{6}x^2 + \tfrac{7}{360}x^4 \ldots.$$

which are convergent when x lies between $+\pi$ and $-\pi$, we may write

$$\coth \mu l = \frac{1}{\mu l} + \frac{\mu l}{3}, \quad \operatorname{cosech} \mu l = \frac{1}{\mu l} - \frac{\mu l}{6}.$$

In both approximations the error is less than the first term omitted, and in our experiments was always less than 0·1 per cent. of the true value. On substituting the above approximate values of coth μl and cosech μl in (4), we get, on introducing the value of μ from (2) and simplifying,

$$\left(\frac{d\theta}{dx}\right)_{x=0} = -\frac{\theta_1-\theta_0}{l} - \frac{phl}{6\lambda q}(\theta_0+2\theta_1-3\theta_s),$$

$$\left(\frac{d\theta}{dx}\right)_{x=l} = -\frac{\theta_1-\theta_0}{l} + \frac{phl}{6\lambda q}(2\theta_0+\theta_1-3\theta_s).$$

If H_0 and H_l be the amounts of heat flowing per second through the sections $x=0$ and $x=l$, respectively, we have

$$H_0 = -\lambda q\left(\frac{d\theta}{dx}\right)_{x=0} = \lambda q\,\frac{\theta_1-\theta_0}{l} + \frac{phl}{6}(\theta_0+2\theta_1-3\theta_s), \qquad (5\text{A})$$

$$H_l = -\lambda q\left(\frac{d\theta}{dx}\right)_{x=l} = \lambda q\,\frac{\theta_1-\theta_0}{l} - \frac{phl}{6}(2\theta_0+\theta_1-3\theta_s). \qquad (5\text{B})$$

The lateral loss of heat ΔH from the rod between the sections is obtained by subtracting (5B) from (5A), *i.e.*,

$$\Delta H = \tfrac{1}{2}\,phl\,(\theta_0+\theta_1-2\theta_s). \qquad (6)$$

All the quantities in (5A) can be directly measured, except λ and h. By carrying out a second experiment in which θ_s is adjusted to a new value, another relation of the form (5A) is obtained. This, together with (5A), can then be solved for λ and h. By means of (6) we can correct for the lateral losses from the heating coil, and the part of the rod projecting beyond the first thermojunction.

It has been assumed in the foregoing theory that λ is independent of the temperature. This is not strictly the case. Since $\theta_1-\theta_0$ is always small, we may suppose that λ is a linear function of the temperature, *i.e.*,

$$\lambda = \lambda_s[1+\alpha(\theta-\theta_s)] = \lambda_s(1+\alpha z), \qquad (7)$$

where α is the temperature coefficient of λ

The fundamental equation (2) may then be written

$$(1+\alpha z)\frac{d^2z}{dx^2} + \alpha\left(\frac{dz}{dx}\right)^2 = \mu_s^2 z. \qquad (8)$$

By solving (8) and carrying out a simple but tedious calculation, the following relation, in which the temperature variation of λ is allowed for, may be obtained:

$$H_0 = \{\lambda q(\theta_1-\theta_0)/l + \tfrac{1}{6}phl(\theta_0+2\theta_1-3\theta_s)\}\{1+\tfrac{1}{2}\alpha(\theta_0+\theta_1-2\theta_s)\}. \qquad (9)$$

If h be negligibly small, (9) reduces to

$$H_0 = \lambda q(\theta_1-\theta_0)\{1+\tfrac{1}{2}\alpha(\theta_0+\theta_1-2\theta_s)\}/l. \qquad (9\text{A})$$

In the present investigation we have not applied the equations (9) and (9A). The accuracy obtainable in the observed temperatures (about 1/200° C.) is not sufficient to justify their use. Further, as α is very small over the range (19° C. to −183° C.) of our observations, the equations (5A) and (5B) are adequate. In the neighbourhood of −252° C. the variation of λ with the temperature for many metals is very rapid, and a correction for this variation must be made if reasonable accuracy is to be obtained. No observer has yet taken it into account.

Experimental Details.

The copper crystal used in this investigation was kindly presented to us by Dr. W. D Coolidge of the General Electric Company, U S A. It is a cylindrical casting, about 12 cm. long and 0·6 cm. in diameter. As the diameter is constant over the entire length to within a few thousandths of an inch, it was decided not to risk machining it.

Two forms of apparatus were used; one for room temperature determinations and another slightly modified form for the lower temperatures. The salient features of the apparatus may be gathered from figs. 1A and 1B. The heating coil, H, had a resistance of 68 ohms. For the room temperature determinations a differential constantan-copper couple was used to determine $(\theta_1 - \theta_0)$, the copper element being the crystal itself. For the other determinations two constantan-copper couples were employed. The junctions were in both cases nearly 10 cm. apart, and were fixed to the crystal by a minute touch of solder The copper water-cooling device, W, was fixed to the glass tube by means of a wax joint. For the low temperature measurements the apparatus (fig. 1A) was inverted, and for W a simple steel cap, M, was substituted The copper stud, S, was hard soldered to the M, and the crystal with portion of the pointed end cut off soldered into the stud. A glass metal joint was made by platinizing the end of the glass tube and then simply soldering the steel cap directly on to it. This type of joint proved quite satisfactory at liquid oxygen temperatures. A high vacuum was obtained by means of the conventional rotary oil pump and a tube containing carefully activated charcoal, which was thoroughly de-gassed before immersion in liquid air. That a high vacuum was obtained was shown by a Pirani gauge.

The materials used for the copper-constantan thermocouples were carefully tested, and only perfectly homogeneous wires were used. The calibration of the differential thermocouples over the range 0° to 30° C., and the two thermocouples over the range −190° C. to 0° C., was effected by direct comparisons of their readings against those of an accurate platinum thermometer, No. 7,

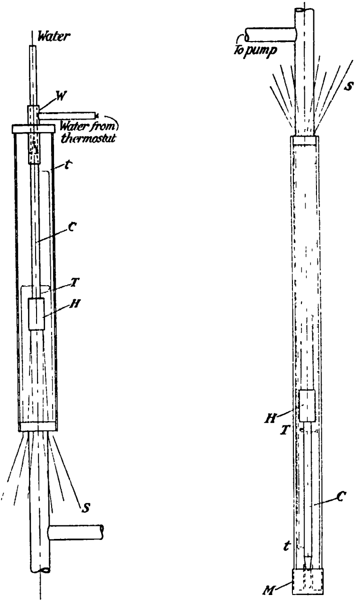


FIG. 1A. FIG. 1B.

C, Crystal. H, Heating Coil. S, Wires to Selector Switch. T, *t*, Thermojunctions
W, Water Cooler. M, Mutal Cap.

used in a previous research by one of us.* For the range — 190° C. to 0° C., No. 7 itself was calibrated by the method recommended by F. Henning.†

---

* See Laby and Hercus, 'Phil. Trans.,' A, vol. 227, p. 63 (1927).

† See F. Henning, 'Temperaturmessung,' p. 101, J. Springer.

We converted our thermocouple readings to temperatures by determining the departure of each couple from the theoretical couple of L. H. Adams.*

All the necessary measurements, the E.M.Fs. of the thermocouples, the current in the heating coil and the drop of potential over its ends, were obtained by means of a Diesselhorst " thermokraftfrei " potentiometer made by Wolf (Berlin). The subsidiary apparatus consisted of 0·01 and 0·1 ohm standard resistances, by Tinsley and Hartmann Braun respectively, for determining the currents in the heating coil. As the drop of potential over the heating coil was too great to measure directly on the potentiometer, a calibrated Hartmann Braun volt box was used to step down the p.d.

The method of eliminating the effect of the lateral heat losses has already been indicated in the section on the theory. Two equations of the form (5A) were obtained from a pair of experiments which differed only in that the temperature of the water circulating in the cooler, W, was adjusted to give a new value of θ_0, say θ'_0, in the second experiment. In this way the heat losses in the two experiments were made different. Both λ and *h* could then be obtained by solving the following equations :—

$$\left.\begin{aligned} \mathrm{H}_0 &= \lambda q\,(\theta_1 - \theta_0)/l + \tfrac{1}{6}phl\,(\theta_0 + 2\theta_1 - 3\theta_s) \\ \mathrm{H}_0' &= \lambda q\,(\theta_1' - \theta_0')/l + \tfrac{1}{6}\,phl\,(\theta_0' + 2\theta_1' - 3\theta_s') \end{aligned}\right\} . \tag{10}$$

The quantity H_0 is not equal to $\mathrm{E} \times \mathrm{I}/\mathrm{J}$

(E volt, the drop of potential over the heating coil, I amp. the current, $J = 4{\cdot}181$ watt/cal. 20° C.),

but is slightly less on account of the lateral loss, $\Delta\mathrm{H}_0$, from the projecting portion of the crystal and the heating coil.

We may write

$$\mathrm{H}_0 = \mathrm{E} \times \mathrm{I}/\mathrm{J} - \Delta\mathrm{H}_0. \tag{11}$$

The lateral loss, $\Delta\mathrm{H}_0$, was approximately allowed for in the following way :—

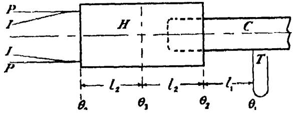


FIG. 2.—H, Heating Coil. P, Potential Leads. I, Current Leads. C, Crystal. T, Thermo-junction.

---

[*] See 'Inter. Crit. Tables,' vol. 1, pp. 57 and 58.

In fig. 2 the maximum temperature of the heating coil, θ_3, is assumed to occur at the middle cross-section, while the temperature of its end sections are supposed each equal to θ_2. If the gradients of temperature in the crystal and coil are inversely proportional to their respective cross-sections Q and q, the loss of heat from the heating coiling and the part of the crystal projecting beyond the thermo-junction θ_1, may be readily shown to be

$$\Delta H_0 = \tfrac{1}{2} phl_1 (\theta_1 + \theta_2 - 2\theta_s) + phl_2 (\theta_2 + \theta_3 - 2\theta_s) + h(2Q - q)(\theta_2 - \theta_s). \tag{12}$$

It was not possible to vary the external conditions in the above manner, so as to allow for the heat losses in the experiments at the lower temperatures. As the values of λ uncorrected for heat losses in the room temperature determination were only 1 per cent. to 2 per cent. greater than the corrected values (see Table I), we concluded that the effect of the lateral losses on λ could be made very small. In the experiments at −183° C. and −79·5° C., the glass tube containing the crystal was immersed in the bath of liquid oxygen or solid carbon dioxide and alcohol to within about 5 cm. of the upper wax joint. The heating coil and thermocouple leads were nowhere in direct contact with the bath, and, in consequence, a small positive residual temperature gradient always occurred in the crystal on account of a small flow of heat down the leads. We allowed for this "*residual*" gradient by plotting the actual differences of temperature, $(\theta_1 - \theta_0)'$, indicated by the thermocouples against the corresponding values of the input of heat $H_0 = E \times I/J$.

It was found that the five points so obtained fitted very accurately along a straight line, with the exception of an intermediate point which was in error by 1/70° C. (see Table II). By extrapolating, the residual gradient (that corresponding to $H_0 = 0$) was obtained. The conductivity λ could then be calculated from the relation

$$H_0 = \lambda q (g - g_0)$$

where g is the observed temperature gradient for the input of energy H_0, and g_0 is the extrapolated value of the gradient for zero input of energy.

The extrapolated value of the residual gradient was always somewhat less than the observed value of the gradient with no current in the heating coil. We were unable to account for this discrepancy satisfactorily.

Results.

(*a*) The details of the room temperature determinations are given below in Table I. The uncorrected values of λ are calculated by means of the simple

formula $H_0 = \lambda q (\theta_1 - \theta_0)/l$; the corrected values by means of equations (10), (11) and (12).

Table I.

E. volts	I amps.	θ_2° C	θ_1° C	θ_0° C	λ uncorrected.	λ corrected.	$h \times 10^5$	Mean temperature.
12 0357 12 0378	0·17660 0·17658	17 6 15·8	28 62 21·29	10 21 2·59	1 016 1·001	0·990	6·9	19 4
12 0365 12 0367	0·19651 0·17655	16·85 15 8	28·64 22·51	10·21 3·80	1·015 1 000	0 988	6·0	19 4
8 0596 8 0587	0·11842 0·11851	18·3 14 9	22·82 10·37	14·147 1·68	1 007 0 968	0 990	4·5	18·6
6 0354 6·0347	0·8878 0·8888	17·4 14·25	20 046 5·64	15 35 0·52	1 0047 0 9226	0·987	4·5	17·7

The mean value of λ between 10·2° and 28·6° C. was

0·989 cal. cm. sec. deg.

with a maximum departure from this mean value of 0·002.

(*b*) From the data of Table I it will be noticed that the uncorrected values of λ do depend on E × I, the watts supplied to the heating coil. For negligibly small losses the values of λ would be independent of E × I.

At the liquid oxygen temperatures, several series of experiments were carried out. By successively improving the high vacuum in the tube, one series was ultimately obtained in which the uncorrected values of λ were practically independent of E × I. For this series then, *h* must have approximated very closely to zero. The details of this series are recorded in Table II.

Table II.

E volts.	I amps.	E × I watts.	θ_1° C.	θ_0° C.	$(\theta_1 - \theta_0)'$.	$(\theta_1 - \theta_0)$.	λ	θ_m.
4·0000	0 06356	0 2542	$-177{\cdot}16_5$	$-179{\cdot}68_1$	$2{\cdot}51_6$	$2{\cdot}04_6$	1·132	−178·42
5·0062	0·07896	0·3954	$-174{\cdot}99_5$	$-178{\cdot}66_2$	$3\ 66_7$	$3{\cdot}19_7$	1·127	−176 83
7·0170	0·10897	0·7647	$-169\ 58_4$	$-176\ 21_2$	$6{\cdot}62_8$	$6{\cdot}15_8$	1·133	−172·90
6·0034	0 09390	0 5637	$-172{\cdot}46_1$	$-177{\cdot}46_3$	$5{\cdot}00_2$	$4{\cdot}53_2$	1·132	−174·96
3·5012	0 05418	0·1897	$-177{\cdot}92_6$	$-179{\cdot}91_8$	$1{\cdot}99_2$	$1\ 52_2$	1·133	−178·92

The quantities in the column $(\theta_1 - \theta_0)'$ are the observed differences of temperature; those in the next column $(\theta_1 - \theta_0)$ are obtained by subtracting

0·470° from each quantity in the column $(\theta_1 - \theta_0)'$. The quantity 0·470° is the extrapolated value of $(\theta_1 - \theta_0)'$ corresponding to $E \times I = 0$.

The values of λ are obtained from the formula

$$E \times I/J = \lambda q (\theta_1 - \theta_0)/l,$$

where $J = 4 \cdot 181$ watts calibrated at 20° C.

The mean value of λ over the range −169·6 to −179·9° C. is

$$1 \cdot 131 \text{ cal. cm. sec. deg.}$$

(*c*) A complete set of experiments with the tube immersed in carbon dioxide was not carried out. The data for two experiments is given in Table III below.

Table III.

E volts.	I amps	E × I watts.	θ_1° C	θ_0° C	$\theta_1 - \theta_0$	λ	θ_m.
2·4842	0·03811	0·0947	$-74 \cdot 04_3$	$-75 \cdot 49_7$	$1 \cdot 45_4$	$1 \cdot 05_4$	−74·77
3·7782	0·05784	0·2185	$-71 \cdot 44_2$	$-73 \cdot 96_0$	$2 \cdot 52_4$		−72·70

To eliminate the effect of the residual gradient in the crystal, the values $(E \times I)$ and $(\theta_1 - \theta_0)$ for the first experiment were subtracted from $(E \times I)$ and $(\theta_1 - \theta_0)'$ for the second experiment, and the conductivity λ calculated from

$$\frac{(E \times I)' - (E \times I)}{J} = \lambda q \left\{ \frac{(\theta_1 - \theta_0)'}{l} - \frac{(\theta_1 - \theta_0)}{l} \right\}.$$

The mean of the values of λ at the mean temperature θ_m of the range over which they were determined are :—

λ	θ_m
	°C.
0·989	19·4
1·054	− 73·74
1·131	−174·8

The only complete series of values of λ for a copper crystal (a naturally occurring one) previously published is that obtained by Schott over the range 0° C. to −252° C. Schott obtained 0·98 cal. cm. sec. deg. for 0° C., which is slightly less than our value $\lambda_{19 \cdot 4^\circ \text{C}} = 0 \cdot 989$. By plotting Schott's values of λ

against the temperature, the value of $\lambda_{-174\cdot8^\circ C.}$ obtained from the curve is roughly 1·25. This is considerably greater than our value of 1·131.

The data obtained both by Schott and by Grüneisen and Goens, as they are not corrected for heat losses, represent, at best, upper limits to the true values of λ. As a result of our experiments we regard it as quite inadequate to rely on a determination of λ for a single value of H_0 at the liquid oxygen and lower temperatures. Unless due precautions are taken to maintain a high vacuum, very large errors in λ are certainly possible.

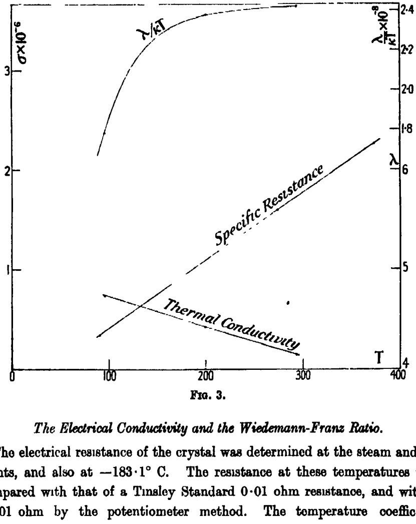


Fig. 3.

The Electrical Conductivity and the Wiedemann-Franz Ratio.

The electrical resistance of the crystal was determined at the steam and ice points, and also at −183·1° C. The resistance at these temperatures was compared with that of a Tinsley Standard 0·01 ohm resistance, and with a 0·001 ohm by the potentiometer method. The temperature coefficient between 0° and 100° C. was found to be equal to that of ordinary electrolytic copper. The specific electrical conductivity κ was not greater than that of ordinary copper, the value of $\kappa_{90\cdot1^\circ K} = 311\cdot6 \times 10^4$ ohm^{-1} cm.$^{-1}$ being appreciably less than that of the purest polycrystal copper. The values we have obtained for the electrical conductivity κ, the thermal conductivity λ, and also the quantity $\lambda/(\kappa T)$ where T is the absolute temperature, are given in the following table. Interpolated values are placed in the parentheses.

Table IV.

Temperature.	κ ohm^{-1} cm.$^{-1}$.	λ watt cm.$^{-1}$ deg.$^{-1}$.	$\lambda/(\kappa T) \times 10^8$.
$37\overset{\circ}{3}\cdot 2$	$44\cdot 4 \times 10^4$	—	—
292·6	(58·4)	4·13	2·42
273·2	$63\cdot 3 \times 10^4$	—	—
199·5	$(92\cdot 7) \times 10^4$	$4\cdot 40_6$	2·38
95·4	$(279\cdot 4) \times 10^4$	$4\cdot 72_9$	1·78
90·1	$311\cdot 6 \times 10^4$	—	—

The values we have obtained for the Wiedemann-Franz constant $\lambda/(\kappa T)$ for $T = 273\cdot 2°$ K. and $T = 90\cdot 1°$ K, and the corresponding values obtained by Schott for the natural crystal used by him, are given below in Table V.

Table V.

Temperature.	$\lambda/(\kappa T) \times 10^8$.	$\lambda/(\kappa T) \times 10^8$ (Schott).
$27\overset{\circ}{3}\cdot 2$ K	2·42	2·42
90·1	1·69	1·80
21	—	1·21

In conclusion, the authors wish to thank Dr. W. D. Coolidge for kindly presenting the copper crystal. The authors are indebted to the Australian Oxygen Company for the liquid oxygen used in the experiments. A Research Scholarship and Grant by the University of Melbourne to one of the authors made the investigation possible.

The General Sampling Distribution of the Multiple Correlation Coefficient.

By R. A. Fisher, Sc.D., Rothamsted Experimental Station, Harpenden, Herts.

(Communicated by Sir James Jeans, Sec R S.—Received August 24, 1928.)

1. *Introductory.*

Of the problems of the exact distribution of statistics in common use that of the multiple correlation coefficient is the last to have resisted solution. It will be seen that the solution introduces an extensive group of distributions, occurring naturally in the most diverse types of statistical investigation, and which in their mere mathematical structure supply an extension in a new direction of the entire system of distributions previously obtained, which in their generality underlie the analysis of variance. The individual distributions of this system were in each case obtained by the exact investigation of a particular problem. It was realised only gradually that many of these distributions, disguised by the different notations appropriate to different problems, were in reality equivalent, and could be made available in practice by a single system of tables. The remaining cases, with the notable exception of the correlation coefficient, then fall into place as particular limiting forms of a single general distribution. As the practical utility of these earlier solutions depends greatly on a recognition of their place in a single system, a very brief account of their mutual relationship may be given.

The only statistic derived from samples of a continuous variate, of which the distribution was known before the present century, appears to be the arithmetic mean of a sample drawn from the normal distribution. In addition, however, two distributions which may be regarded as distributions of statistics had also been found, namely, Bernoulli's binominal distribution, and Poisson's series. Both of these distributions possess the property that the aggregate of the values of a sample is itself distributed in a distribution of the same type. In all three classical cases, therefore, the distribution of the statistic derived from a finite sample was known only by a mathematical simplification of this special type. In all other cases, approximations of unknown accuracy based on the use of the standard error and the assumption of normal distributions had perforce to be used.

In 1908 "Student"[*] attacked the problem of the distribution of the mean of a normal sample measured, as in practice it must usually be, in terms of the standard error as estimated from the same sample. He was thus incidentally led to the equally fundamental distribution of the variance of a normal sample. This latter, to which the general distribution of the analysis of variance degenerates when n_2 tends to infinity, is in reality equivalent to the distribution found by Pearson[†] in 1900 for the χ^2 measure of discrepancy developed for testing goodness of fit. From this "Student" was able to derive the exact distribution (the distribution of t) of the mean of a unique sample, which as subsequently appeared falls into the same system with $n_1 = 1$. The two principal limiting forms of the general distribution were thus known in 1908, and were available for practical application by means of Elderton's[‡] and "Student's"[*] first tables.

In 1915 the distribution of the coefficient of correlation was obtained[§] by a use of Euclidean hyper-space similar to that employed below. The same method served at the same time to put "Student's" results upon a rigorous basis. The distribution of the correlation coefficient stands outside the analysis of variance system, but, as will be seen in the present paper, it is brought into coherent connection with it by the distribution of the coefficient of multiple correlation. When, however, the corresponding distribution of the intraclass correlation was obtained[||], the distribution found was of a new and different type, which, as subsequently appeared, was the general distribution of the analysis of variance, in which the variance is analysed into two parts representing that within and that between the classes or "fraternities" of which the data are composed. This was the first instance of the general distribution which from the notation there used is distinguished as the distribution of z.

The recognition of the fundamental importance of the two parameters, n_1 and n_2, which specify the numbers of degrees of freedom in the two estimates of variance to be compared, and the recognition of the distribution of χ^2 as equivalent to that of an estimate of variance led, in 1922 and the following two years,[¶] to the demonstration that it is always the number of degrees of freedom

---

[*] 'Biometrika,' vol. 6, p. 1 (1908).

[†] 'Phil. Mag.,' vol. 50, p. 157 (1900).

[‡] 'Biometrika,' vol. 1, p. 155 (1902).

[§] Fisher, 'Biometrika,' vol. 10, p. 507 (1915).

[||] 'Metron.,' vol. 1, No. 4, p. 1 (1921).

[¶] 'J. R. Stat. Soc.,' vol. 85, p. 87 (1922); 'Economica,' vol. 3, p. 139 (1923), 'J. R. Stat. Soc.,' vol. 87, p. 442 (1924).

which is to be used in applying the test of goodness of fit. The further proof that the test is only valid when the methods of estimation employed have been *efficient*, binds the theory of goodness of fit closely to that or estimation in the development of which the exact distribution of statistics play an essential part.[*]

Meanwhile,[†] a solution of the exact distribution of χ^2 when applied to test the goodness of fit of regression formulæ had shown that a modification was required in this case, which, in fact, involved dropping the approximate assumption that n_2 was infinite, and reduced the general distribution to the same form as that already found in the study of intraclass correlation. At the same time, the distribution of the correlation ratio, η, derived from uncorrelated material, was shown to belong to the same class with n_1, equal to one less than the number of arrays; and the distribution of regression coefficients, whether total or partial, and whether employed in a linear or a non-linear formula, were shown to conform to "Student's" distribution. The solution of the distribution of the correlation ratio η really included also that of the multiple correlation coefficient for samples drawn from uncorrelated material, the distribution of which was given in its appropriate notation in 1924.[‡]

In the same year[§] it was found possible to use the representation in hyperspace to demonstrate that the distribution of the partial correlation coefficients is exactly the same as that primarily found for the total correlation, provided that unity is deducted from the sample number for each variate eliminated.

Each distinct type of distribution found has thus occurred repeatedly in different investigations; whereas, however, nearly all cases are reducible to a common type capable of exact treatment by the same simple arithmetical procedure,[‖] and requiring the same fundamental table, the distribution of the (intraclass) correlation coefficient, total or partial, stood aside from the main system, and was capable of only an approximate treatment by using the distribution of z.

The distribution of the multiple correlation coefficient, apart from the practical necessity of applying to observed results sufficiently exact tests of significance, is thus of great theoretical interest owing to the close connection which must exist between it and the simple correlation coefficient, on the one hand, and, on the other, to the form already obtained from uncorrelated material.

---

[*] 'Phil. Trans.,' A, vol. 222, p. 309 (1922).

[†] 'J. R. Stat. Soc.,' vol. 85, p. 597 (1922).

[‡] 'Phil. Trans.,' B, vol. 213, p. 89 (1924).

[§] 'Metron.,' vol. 3, p. 329 (1924).

[‖] 'Statistical Methods for Research Workers,' 2nd ed., Oliver & Boyd, Edinburgh, 1928.

The latter solution involves, besides the variate and frequency, the two parameters n_1 and n_2, and is therefore a functional relation between four variables. The new solution necessarily involves also the multiple correlation in the population sampled, making a fifth variable; complete tabulation of the results would thus require a table of fourfold entry; even confining attention to specified points of special importance, such as the 5 per cent. and 1 per cent. points, a procedure that has made tabulation practicable for the distribution of z, we should still have tables of triple entry. The problem of adequate tabulation is certainly not insurmountable, but to ascertain the proper method to adopt in its presentation will require further study of the nature of the function. The table of the 5 per cent. points of the distribution of B (Section 5) should in the author's opinion provide sufficient guidance for the greater number of practical applications.

2. *Method of Solution.*

The primary problem of the sampling distribution of the correlation coefficient between two variates, x and y, was originally solved by interpreting the n individual values of either variate appearing in the sample as the co-ordinates of a point in Euclidean space of n dimensions. It then easily appeared that the correlation coefficient between the variates was the cosine of the angle between the two *radii vectores* drawn from the origin to points, the co-ordinates of which represented the deviations from the mean of the sample of the two variates concerned.

The frequency with which r, the observed correlation coefficient, falls in any infinitesimal range dr may be usefully thought of as the product of two factors, one being the generalised volume in which the second sample point may lie so that the correlation may fall within the assigned range, this value being independent of the correlational properties of the population sampled, while the second is a factor by which the frequency density in any element of volume is modified by the correlation between x and y in the population. With zero correlation in the population, the frequency density at any point depends only on its distance from the origin, and since for any given distance the point is free to move over a sphere in $(n-1)$ dimensions, one dimension having been eliminated by using the sample mean as origin, it is easy to see that for this case the frequency distribution of r is given by

$$df = \frac{[\frac{1}{2}(n-3)]!}{[\frac{1}{2}(n-4)]!\sqrt{\pi}}(1-r^2)^{\frac{1}{2}(n-4)}\,dr.$$

The general solution of the primary problem*

$$df = \frac{n-2}{\pi}(1-\rho^2)^{\frac{1}{2}(n-1)}(1-r^2)^{\frac{1}{2}(n-4)}\int_0^\infty \frac{dz}{(\cosh z - \rho r)^{n-1}} \, . \, dr,$$

may be written with advantage

$$df = \frac{[\frac{1}{2}(n-3)]\,!}{[\frac{1}{2}(n-4)]\,!\,\sqrt{\pi}}(1-r^2)^{\frac{1}{2}(n-4)}\,dr$$
$$\times \frac{[\frac{1}{2}(n-2)]\,!}{[\frac{1}{2}(n-3)]\,!\,\sqrt{\pi}}(1-\rho^2)^{\frac{1}{2}(n-1)}\int_{-\infty}^\infty \frac{dz}{(\cosh z - \rho r)^{n-1}} \, .$$

The second factor then represents the effect upon the frequency density, in the region represented by *dr*, of a correlation ρ in the sampled population: the numerical part of this factor is merely such as to reduce it to unity when $\rho = 0$.

With multiple correlations we are concerned with the correlations between a dependent variate *y*, and a number of independent variates, $x_1, x_2, \ldots, x_{n_1}$, and, moreover, with the correlations of the latter among themselves. It was not at first obvious that the sampling distribution did not involve this whole matrix of correlations, in which case, even if it could be determined, it would be of no practical use. The argument, by which it can be seen to depend from only a single parameter of the population, is therefore of special interest, as by its general character it is applicable to a number of statistical problems, and leads in this case directly to the solution.

The multiple correlation of *y* with $x_1, x_2, \ldots, x_{n_1}$ is the correlation between *y* and that linear function of $x_1, x_2, \ldots, x_{n_1}$ with which its correlation is highest. If, therefore, for the dependent variates, *x*, we substitute an equal number of new variates, ξ, defined as linear functions of the n_1 variates, *x*, then the multiple correlation in the population, and in every sample from it, will remain unchanged. In particular we may choose as ξ_1, that linear function the correlation of which with *y* in the population sampled is highest, and for the remaining variates, ξ, we can choose linear functions of *x*, uncorrelated with ξ_1, or with each other. In choosing the last of these we have no more than $n_1 - 1$ conditions to be satisfied by the ratios of n_1 coefficients. If this is done it is easy to see, or to demonstrate, that all of the variates ξ, except ξ_1, have zero correlation with *y*. Using the variates ξ the sampling distribution of the multiple correlation R can only depend on the correlation in the population sampled between ξ_1 and *y*, namely, on the multiple correlation in the population sampled, which we may designate by ρ.

---

* Fisher, 'Biometrika,' vol. 10, p. 507 (1915).

The geometrical interpretation of the multiple correlation coefficient R is that it is the cosine of the angle between the *radius vector* of the dependent variate y and the planar region including the *radii vectores* of the n_1 independent variates; its distribution when $\rho = 0$, which depends only on the geometrical elements of volume, has been thus shown* to be

$$df = \frac{[\frac{1}{2}(n_1 + n_2 - 2)]\,!}{[\frac{1}{2}(n_1 - 2)]\,!\,[\frac{1}{2}(n_2 - 2)]\,!}(\mathrm{R}^2)^{\frac{1}{2}(n_1-2)}(1 - \mathrm{R}^2)^{\frac{1}{2}(n_2-2)}\,d(\mathrm{R}^2),$$

where n, the sample number, is replaced by $n_1 + n_2 + 1$; but in what way this distribution is modified when ρ is not zero has been hitherto entirely unknown.

It is evident, however, that since we have reduced the problem of the multiple correlation coefficient to one involving only a single correlation, the frequency density of any configuration will be affected merely by a factor

$$\frac{[\frac{1}{2}(n-2)]\,!}{[\frac{1}{2}(n-3)]\,!\,\sqrt{\pi}}(1 - \rho^2)^{\frac{1}{2}(n-1)}\int_{-\infty}^{\infty}\frac{dz}{(\cosh z - \rho r)^{n-1}},$$

in which r is the correlation in the sample between y and ξ_1; this factor will, however, vary, because r varies in the different configurations which give rise to the same value of R. Consider now a third variate, Y, representing the linear function of the independent variates which in the sample is most closely correlated with y, or, in other words, the prediction formula for y Its correlation with ξ_1 we may represent by $\cos\psi$, and since the partial correlation of y with ξ_1 (or any other linear function of the independent variates) when Y is eliminated, must be zero, it is evident that

$$r = \mathrm{R}\cos\psi.$$

For a given value of ψ, therefore, the density factor is constant in the different configurations which give the same value of R, but, in the absence of correlation, the frequency with which ψ falls in the range $d\psi$ is evidently

$$\frac{[\frac{1}{2}(n_1 - 2)]\,!}{[\frac{1}{2}(n_1 - 3)]\,!\,\sqrt{\pi}}\sin^{n_1-2}\psi\,d\psi;$$

hence integrating over all values of ψ, the density factor becomes

$$\frac{(1-\rho^2)^{\frac{1}{2}(n_1+n_2)}}{\pi}\cdot\frac{[\frac{1}{2}(n_1 + n_2 - 1)]\,!}{[\frac{1}{2}(n_1 + n_2 - 2)]\,!}$$
$$\times\frac{[\frac{1}{2}(n_1 - 2)]\,!}{[\frac{1}{2}(n_1 - 3)]\,!}\int_0^{\pi}d\psi\int_{-\infty}^{\infty}\frac{\sin^{n_1-2}\psi\,.\,dz}{(\cosh z - \rho\mathrm{R}\cos\psi)^{n_1+n_2}},$$

---

* Fisher, 'Phil. Trans.,' B, vol. 213, p. 89 (1924).

and the complete expression for the distribution of R is

$$df = \frac{[\frac{1}{2}(n_1+n_2-1)]!}{[\frac{1}{2}(n_2-2)]!\,[\frac{1}{2}(n_1-3)]!}\cdot\frac{(1-\rho^2)^{\frac{1}{2}(n_1+n_2)}}{\pi}$$
$$\times\,(\mathrm{R}^2)^{\frac{1}{2}(n_1-2)}(1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}\,d(\mathrm{R}^2)\int_0^\pi d\psi\int_{-\infty}^{\infty}\frac{\sin^{n_1-2}\psi\,.\,dz}{(\cosh z-\rho\mathrm{R}\cos\psi)^{n_1+n_2}}.$$

3. *The Hypergeometric Form.*

Apart from the factor,

$$(1-\rho^2)^{\frac{1}{2}(n_1+n_2)},$$

the density factor may be reduced to a hypergeometric function. For in the expression,

$$\frac{1}{\pi}\cdot\frac{[\frac{1}{2}(n_1+n_2-1)]!}{[\frac{1}{2}(n_1+n_2-2)]!}\cdot\frac{[\frac{1}{2}(n_1-2)]!}{[\frac{1}{2}(n_1-3)]!}\int_0^\pi d\psi\int_{-\infty}^{\infty}\frac{\sin^{n_1-2}\psi\,.\,dz}{(\cosh z-\rho\mathrm{R}\cos\psi)^{n_1+n_2}},$$

the integrand may be expanded in the uniformly convergent series

$$\sum_{t=0}^{\infty}\frac{(n_1+n_2+2t-1)!}{(n_1+n_2-1)!\,(2t)!}\cdot\frac{\cos^{2t}\psi\sin^{n_1-2}\psi}{\cosh^{n_1+n_2+2t}z}(\rho^2\mathrm{R}^2)^t,$$

in which the odd powers of cos ψ, which evidently disappear on integration, have been omitted. Remembering now that

$$\int_0^\pi\cos^{2t}\psi\sin^{n_1-2}\psi\,d\psi=\frac{[\frac{1}{2}(2t-1)]!\,[\frac{1}{2}(n_1-3)]!}{[\frac{1}{2}(n_1+2t-2)]!},$$

and

$$\int_{-\infty}^{\infty}\frac{dz}{\cosh^{n_1+n_2+2t}z}=\frac{[\frac{1}{2}(n_1+n_2+2t-2)]!\sqrt{\pi}}{[\frac{1}{2}(n_1+n_2+2t-1)]!},$$

we have a power series for the integral, which may be written

$$\frac{[\frac{1}{2}(n_1-2)]!}{[\frac{1}{2}(n_1+n_2-2)]!^2}\sum_{t=0}^{\infty}\frac{[\frac{1}{2}(n_1+n_2+2t-2)]!^2}{t!\,[\frac{1}{2}(n_1+2t-2)]!}(\rho^2\mathrm{R}^2)^t,$$

or

$$\mathrm{F}[\tfrac{1}{2}(n_1+n_2),\ \ \tfrac{1}{2}(n_1+n_2),\ \ \tfrac{1}{2}n_1,\ \rho^2\mathrm{R}^2],$$

so that the distribution of R obtained in section 2 takes the form

$$df=\frac{[\frac{1}{2}(n_1+n_2-2)]!}{[\frac{1}{2}(n_1-2)]!\,[\frac{1}{2}(n_2-2)]!}(1-\rho^2)^{\frac{1}{2}(n_1+n_2)}\mathrm{F}[\tfrac{1}{2}(n_1+n_2),\tfrac{1}{2}(n_1+n_2),\tfrac{1}{2}n_1,\rho^2\mathrm{R}^2]$$
$$\times\,(\mathrm{R}^2)^{\frac{1}{2}(n_1-2)}(1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}\,d(\mathrm{R}^2).\quad\text{(A)}$$

4. *Elementary Cases.*

4.1. ***When n_2 is even.***—When n_2 is even the identity,

$$F\left[\tfrac{1}{2}(n_1+n_2),\ \tfrac{1}{2}(n_1+n_2),\ \tfrac{1}{2}n_1,\ \rho^2R^2\right]$$
$$= (1-\rho^2R^2)^{-\frac{1}{2}(n_1+2n_2)} - F\left(-\tfrac{1}{2}n_2,\ -\tfrac{1}{2}n_2,\ \tfrac{1}{2}n_1,\ \rho^2R^2\right),$$

gives a terminating series.

Thus, when $n_2 = 2$, we have the series of distributions

$$df = (1-\rho^2)^{\frac{1}{2}(n_1+2)}(1-\rho^2R^2)^{-\frac{1}{2}(n_1+4)}(n_1+2\rho^2R^2)\,R^{n_1-1}\,dR,$$

having the special forms

$$(2.2) \qquad df = (1-\rho^2)^2(2+2\rho^2R^2)(1-\rho^2R^2)^3\,.\,R\,dR,$$

$$(3.2) \qquad df = (1-\rho^2)^{5/2}(3+2\rho^2R^2)/(1-\rho^2R^2)^{7/2}\,.\,R^2\,dR,$$

$$(4.2) \qquad df = (1-\rho^2)^3(4+2\rho^2R^2)/(1-\rho^2R^2)^4\,.\,R^3\,dR$$

when n_1 is 2, 3 and 4.

When $n_2 = 4$, we have a somewhat less simple series of distributions

$$df = \frac{2(1-\rho^2)^{\frac{1}{2}(n_1+4)}}{(1-\rho^2R^2)^{\frac{1}{2}(n_1+8)}}\left\{\frac{n_1(n_1+2)}{2\,.\,4}+2\,\frac{n_1+2}{2}\,\rho^2R^2+\rho^4R^4\right\}R^{n_1-2}(1-R^2)\,d(R^2),$$

and when $n_2 = 6$, a series which may be written

$$df = \frac{3(1-\rho^2)^{\frac{1}{2}(n_1+6)}}{(1-\rho^2R^2)^{\frac{1}{2}(n_1+12)}}\left\{\frac{n_1(n_1+2)(n_1+4)}{2\,.\,4\,.\,6}+3\,\frac{(n_1+2)(n_1+4)}{2\,.\,4}\,\rho^2R^2\right.$$
$$\left.+3\,\frac{n_1+4}{2}\,\rho^4R^4+\rho^6R^6\right\}R^{n_1-2}(1-R^2)^2\,d(R^2),$$

an expression in which the general method of formation of the terms is readily seen.

4.2. ***When n_1 and n_2 are both odd.***—A second group of cases in which the frequency element is expressible in finite terms in elementary functions occurs when both n_1 and n_2 are odd. If, for example, we put $n_1 = 3$ in the expression

$$\int_0^\pi d\psi \int_{-\infty}^{\infty} \frac{\sin^{n_1-2}\psi\,.\,dz}{(\cosh z - \rho R\cos\psi)^{n_1+n_2}},$$

and integrate with respect to $\cos\psi$, we obtain

$$\int_{-\infty}^{\infty} \frac{dz}{(n_2+2)\,\rho R}\left\{(\cosh z-\rho R)^{-(n_2+2)} - (\cosh z+\rho R)^{-n_2+2}\right\},$$

a form of integral which, as was shown in the case of the simple correlation coefficient*, is expressible in finite terms by the aid of the circular functions.

---

* Fisher, 'Biometrika,' vol. 10, p. 507 (1915).

For

$$\int_{-\infty}^{\infty} \frac{dz}{\cosh z - \rho R} = \frac{2\theta}{\sin\theta},$$

where $\cos\theta = -\rho R$, and θ does not exceed the bounds 0 to π; hence

$$\int_{-\infty}^{\infty} \frac{dz}{(\cosh z - \rho R)^{n_2+2}} = \frac{2}{(n_2+1)!}\left(\frac{d}{d\cos\theta}\right)^{n_2+1} \frac{\theta}{\sin\theta},$$

and therefore, if n_2 is odd,

$$\frac{1}{(n_2+2)\,\rho R}\int_{-\infty}^{\infty} \{(\cosh z - \rho R)^{-(n_2+2)} - (\cosh z + \rho R)^{-(n_2+2)}\}\,dz$$
$$= \frac{4}{(n_2+2)!\,\rho R}\left(\frac{d}{d(\rho R)}\right)^{n_2+1} \frac{\sin^{-1}\rho R}{\sqrt{1-\rho^2R^2}}.$$

Hence for the determination of the simpler distributions of this series we require

$$\left(\frac{d}{\cos\phi\,d\phi}\right)^2 \frac{\phi}{\cos\phi} = \frac{1}{\cos^3\phi}(\phi + 3\tan\phi + 3\phi\tan^2\phi)$$

$$\left(\frac{d}{\cos\phi\,d\phi}\right)^4 \frac{\phi}{\cos\phi} = \frac{1}{\cos^5\phi}(9\phi + 55t + 90\phi t^2 + 105t^3 + 105\phi t^4)$$

$$\left(\frac{d}{\cos\phi\,d\phi}\right)^6 \frac{\phi}{\cos\phi} = \frac{9}{\cos^7\phi}(25\phi + 231t + 525\phi t^2 + 1190t^3$$
$$+ 1575\phi t^4 + 1155t^5 + 1155\phi t^6,$$

which lead directly to the distributions,

$$(3.1)\qquad df = \frac{1}{\pi}(1-\rho^2)^2\,(1-R^2)^{-\frac{1}{2}}\,(1-\rho^2R^2)^{-2}\,\{3 + \alpha\,(1+2\rho^2R^2)\}\,R^2dR,$$

in which α stands for

$$\frac{\sin^{-1}(\rho R)}{\rho R\sqrt{1-p^2R^2}} = 1 + \tfrac{2}{3}\rho^2R^2 + \frac{2\,.\,4}{3\,.\,5}\rho^4R^4 + \quad \ldots\ \textit{ad inf.}$$

$$(3\ 3)\qquad df = \frac{1}{4\pi}(1-\rho^2)^3\,(1-R^2)^{\frac{1}{2}}\,(1-\rho^2R^2)^{-4}\,R^2dR$$
$$\times\{5\,(11 + 10\rho^2R^2) + 3\alpha\,(3 + 24\rho^2R^2 + 8\rho^4R^4)\},$$

$$(3\ 5)\qquad df = \frac{1}{8\pi}(1-\rho^2)^4\,(1-R^2)^{3/2}\,(1-\rho^2R^2)^{-6}\,R^2\,dR$$
$$\times\{7\,(33 + 104\rho^2R^2 + 28\rho^4R^4) + 5d\,(5 + 90\rho^2R^2 + 120\rho^4R^4 + 16\rho^6R^6)\}.$$

A similar process of integration is available for other odd values of n_1; for $n_1 = 5$ we have the distributions

$$(5.1)\qquad df = \frac{1}{4\pi\rho^2}(1-\rho^2)^3(1-R^2)^{-\frac{1}{2}}(1-\rho^2R^2)^{-3}R^2\,dR$$
$$\times\{1+14\rho^2R^2+\alpha(-1+8\rho^2R^2+8\rho^4R^4)\}.$$

$$(5.3)\qquad df = \frac{3}{8\pi\rho^2}(1-\rho^2)^4(1-R^2)^{\frac{1}{2}}(1-\rho^2R^2)^{-5}R^2\,dR$$
$$\times\{1+68\rho^2R^2+36\rho^4R^4+\alpha(-1+18\rho^2R^2+72\rho^4R^4+16\rho^6R^6)\}.$$

$$df = \frac{1}{64\pi\rho^2}(1-\rho^2)^5(1-R^2)^{3/2}(1-\rho^2R^2)^{-7}R^2\,dR$$
$$(5.5)\qquad \times\{25+4678\rho^2R^2+8664\rho^4R^4+1648\rho^6R^6$$
$$+\alpha(-25+800\rho^2R^2+7200\rho^4R^4+6400\rho^6R^6+640\rho^8R^8)\}.$$

The polynomial coefficient of α in the hypergeometric function is itself easily expressed in terms of a function of this sort, in the forms

$$\frac{3\,.\,5\ldots(n_1-2)\,.\,3^2\,.\,5^2\ldots n_2^2}{2^2\,.\,4^2\ldots(n_1+n_2-2)^2}(-\rho^2R^2)^{-\frac{1}{2}(n_1-3)}\,F[-\tfrac{1}{2}(n_1+n_2-2),\ \tfrac{1}{2}(n_1+n_2-2),$$
$$\tfrac{1}{2}(2-n_1),\ \rho^2R^2],$$

or

$$\frac{n_1-2}{2\,.\,4\ldots(n_1+n_2-2)}(\rho^2R^2)^{\frac{1}{2}(n_2+1)}\,F\left[-\frac{n_2}{2},\ -\frac{n_1+n_2-2}{2},\ 1,\ \frac{1}{\rho^2R^2}\right];$$

from this the remainder may in any particular case be found fairly easily by equating coefficients in the initial terms of the expansion of

$$F(\tfrac{1}{2}(n_1+n_2),\tfrac{1}{2}(n_1+n_2),\tfrac{1}{2}n_1,\rho^2R^2).$$

5. *The Problem of Large Samples.*

Some confusion has been caused by the fact that, while for any finite value of ρ, however small, the distribution of R will be normal for a sufficiently large sample, yet when $\rho = 0$ the distribution is far from normal. The approximate distribution appropriate to the theory of large samples, for different values of $\rho\sqrt{n_2}$, may be found as follows.

If we write $n_2\rho^2 = \beta^2$, $n_2R^2 = B^2$, and allow n_2 to increase indefinitely, the limiting form taken by the general distribution is

$$df = \frac{(\frac{1}{2}B^2)^{\frac{1}{2}(n_1-2)}}{[\frac{1}{2}(n_1-2)]\,!}e^{-\frac{1}{2}B^2-\frac{1}{2}\beta^2}\left\{1+\frac{1}{n_1}\frac{\beta^2B^2}{2}+\frac{1}{n_1(n_1+2)}\frac{\beta^2B^2}{2\,.\,4}+\ldots\right\}d(\tfrac{1}{2}B^2),\ \text{(B)}$$

which may be written in terms of a Bessel function as

$$(B/i\beta)^{\frac{1}{2}(n_1-2)}e^{-\frac{1}{2}(B^2-\beta^2)}\,.\,J_{\frac{1}{2}(n_1-2)}(i\beta B)\,.\,d(\tfrac{1}{2}B^2).$$

When n_1 is odd, these may be reduced to elementary functions; thus for $n_1 = 3$, we have

$$df = \frac{1}{\sqrt{2\pi}}\frac{B}{\beta}\left\{e^{-\frac{1}{2}(B-\beta)^2} - e^{-\frac{1}{2}(B+\beta)^2}\right\}.\,dB,$$

an interesting distribution which connects the extreme forms found by making β zero for uncorrelated populations, and large for populations with a significant though still small correlation. When $\beta = 0$, we have

$$df = (2/\pi)^{\frac{1}{2}}\,B^2 \exp\left(-\tfrac{1}{2}B^2\right) dB,$$

the distribution of x for 3 degrees of freedom, while when β is large, B is distributed normally about β, in the form

$$df = (2\pi)^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}(B-\beta)^2\right\} dB,$$

and therefore R is distributed normally about ρ, with variance which may be equated to $1/n_2$.

When $n_1 = 5$, the system of distributions is

$$df = \frac{1}{\sqrt{2\pi}}\frac{B^2}{\beta^2}\left\{\left(1-\frac{1}{\beta B}\right)\exp\left\{-\tfrac{1}{2}(B-\beta)^2\right\} + \left(1+\frac{1}{\beta B}\right)\exp\left\{-\tfrac{1}{2}(B+\beta)^2\right\}\right\} dB,$$

and when $n_1 = 7$

$$df = \frac{1}{\sqrt{2\pi}}\frac{B^3}{\beta^3}\left\{\left(1-\frac{3}{\beta B}+\frac{3}{\beta^2 B^2}\right)\exp\left\{-\tfrac{1}{2}(B-\beta)^2\right\} - \left(1+\frac{3}{\beta B}+\frac{3}{\beta^2 B^2}\right)\exp\left\{-\tfrac{1}{2}(B+\beta)^2\right\}\right\} dB,$$

In the cases in which n_1 is even, the probability of exceeding a given value B may be written

$$\int_B^{\infty} 2^{-\frac{1}{2}(n_1-2)}\, e^{-\frac{1}{2}\beta^2} \sum_{t=0}^{\infty} \frac{\beta^{2t}}{2^{2t}\,.\,t\,!\,\left[\frac{1}{2}(n_1+2t-2)\right]!}\, x^{n_1+2t-1}\, e^{-\frac{1}{2}x^2}\, dx\,;$$

using the fact that when k is odd

$$\int_B^{\infty} \frac{x^k}{2^{\frac{1}{2}(k-1)}.\left[\frac{1}{2}(k-1)\right]!}\, e^{-\frac{1}{2}x^2}\, dx = e^{-\frac{1}{2}B^2}\left\{1+\frac{B^2}{2}+\frac{B^4}{2\,.\,4}+\ldots+\frac{B^{k-1}}{2\,.\,4\ldots(k-1)}\right\},$$

the integral becomes

$$\sum_{t=0}^{\infty} e^{-\frac{1}{2}\beta^2}\frac{\left(\frac{1}{2}\beta^2\right)^t}{t\,!} \sum_{u=0}^{t+\frac{1}{2}(n_1-2)} e^{-\frac{1}{2}B^2}\frac{\left(\frac{1}{2}B^2\right)^u}{u\,!},$$

involving only the terms of two Poisson Series with mean values $\frac{1}{2}\beta^2$ and $\frac{1}{2}B^2$. If t and u be regarded as variates distributed independently in two such series, the probability may be identified with the probability that u should not exceed t by $\frac{1}{2}n_1$, or more.

The distributions developed in this section are limiting forms appropriate to large samples, in which exact account is taken of the positive bias of small observed multiple correlations; they will provide at least an approximate treatment of those cases of great practical importance in which n_2 does not exceed 100, and in which, therefore, the positive bias is prominent for observed values of R which are not small. The fact that sampling errors of the simple correlation coefficient have been successfully represented by a normal distribution by means of the transformation $z = \tanh^{-1} r$, suggests that pending fuller tests than are at present practicable, the transformation

$$\mathrm{B} = \sqrt{n_2} \tanh^{-1} \mathrm{R}, \quad \beta = \sqrt{n_2} \tanh^{-1} \rho,$$

will supply tests of significance of precision, sufficient for practical purposes, in the important region alluded to.

Table I (table of B) shows the 5 per cent. points of these distributions, for

Table of 5 per cent. points of the distribution of B.

Values of β_1	Value of n_1						
	1.	2.	3.	4.	5.	6.	7.
0	1·9600	2·4477	2·7955	3·0802	3·3272	3·5485	3·7506
0·2	1·9985	2·4720	2·8140	3·0955	3·3405	3·5602	3·7613
0·4	2·1070	2·5419	2·8680	3·1405	3·3796	3·5951	3·7930
0·6	2·2654	2·6497	2·9533	3·2125	3·4426	3·6517	3·8445
0·8	2·4505	2·7855	3·0640	3·3076	3·5268	3·7278	3·9144
1·0	2·6461	2·9398	3·1941	3·4216	3·6291	3·8210	4·0005
1·2	2·8451	3·1059	3·3386	3·5505	3·7462	3·9289	4·1008
1·4	3·0449	3·2796	3·4935	3·6911	3·8756	4·0491	4·2134
1·6	3·2449	3·4584	4·6561	3·8408	4·0148	4·1796	4·3363
1·8	3·4449	3·6410	3·8246	3·9978	4·1620	4·3184	4·4681
2·0	3·6449	3·8263	3·9976	4·1604	4·3158	4·4645	4·6074
2·2	3·8449	4·0137	4·1743	4·3278	4·4750	4·6166	4·7531
2·4	4·0449	4·2027	4·3539	4·4990	4·6388	4·7738	4·9043
2·6	4·2449	4·3932	4·5359	4·6735	4·8065	4·9353	5·0603
2·8	4·4449	4·5847	4·7199	4·8506	4·9774	5·1006	5·2204
3·0	4·6449	4·7772	4·9055	5·0301	5·1512	5·2691	5·3840
3·2	4·8449	4·9705	5·0926	5·2115	5·3273	5·4404	5·5508
3·4	5·0449	5·1644	5·2809	5·3946	5·5056	5·6142	5·7204
3·6	5·2449	5·3589	5·4703	5·5792	5·6857	5·7901	5·8924
3·8	5·4449	5·4914	5·6606	5·7650	5·8675	5·9679	6·0665
4·0	5·6449	5·7493	5·8516	5·9521	6·0506	6·1475	6·2426
4·2	5·8449	5·9451	6·0434	6·1401	6·2351	6·3285	6·4204
4·4	6·0449	6·1412	6·2359	6·3290	6·4206	6·5109	6·5998
4·6	6·2449	6·3376	6·4288	6·5187	6·6072	6·6945	6·7805
4·8	6·4449	6·5342	6·6223	6·7091	6·7947	6·8792	6·9625
5·0	6·6449	6·7311	6·8162	6·9002	6·9831	7·0649	7·1457

values of β from 0 to 5 and of n_1 from 1 to 7. The values tabulated are the values of B which will be exceeded by chance in 5 per cent. random trials, and which therefore give a presumption that β is really greater than the value postulated. Thus, when $n_1 = 3$, it may be seen at a glance that a value $B = 5 \cdot 7$ indicates that β probably exceeds $3 \cdot 8$.

For a great part of the labour of constructing this Table I am indebted to Mr. A. J. Page, I.C.S., whose assistance in my laboratory while on leave has thus enabled me to press forward with the theoretical investigation of the new distributions.

6. *The Probability Integral.*

For calculations involving finite probabilities of occurrence, including tests whether an observed R is or is not significantly discrepant from a hypothetical ρ, it is not the frequency element but its integral that is required. It is fortunate that the frequency distribution we have found when n_2 is even leads to a probability integral of a tolerably simple form.

The frequency element

$$(1-\rho^2)^{\frac{1}{2}(n_1+n_2)} \, \frac{n_1+n_2-2}{2}\,! \cdot \frac{(\mathrm{R}^2)^{\frac{1}{2}(n_1-2)}}{[\frac{1}{2}(n_1-2)]\,!}$$
$$\times \frac{(1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}}{[\frac{1}{2}(n_2-2)]\,!} \, \mathrm{F}\left[\frac{n_1+n_2}{2}, \, \frac{n_1+n_2}{2}, \, \frac{n_1}{2}, \, \rho^2\mathrm{R}^2\right] d\,(\mathrm{R}^2);$$

may be written

$$(1-\rho^2)^{\frac{1}{2}(n_1+n_2)} \sum_{t=0}^{\infty} \frac{[\frac{1}{2}(n_1+n_2+2t-2)]\,!^2}{[\frac{1}{2}(n_1+n_2-2)]\,!\,t\,!} \, \rho^{2t} \, \frac{(\mathrm{R}^2)^{\frac{1}{2}(n_1+2t-2)}}{[\frac{1}{2}(n_1+2t-2)]\,!}$$
$$\times \frac{(1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}}{[\frac{1}{2}(n_1-2)]\,!} \, d\,(\mathrm{R}^2);$$

but if n_2 is even

$$\int_0^{\mathrm{R}^2} \frac{n_1+n_2+2t-2}{2}\,! \, \frac{(\mathrm{R}^2)^{\frac{1}{2}(n_1+2t-2)}}{[\frac{1}{2}(n_1+2t-2)]\,!} \cdot \frac{(1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}}{[\frac{1}{2}(n_2-2)]\,!} \, d\,(\mathrm{R}^2)$$

is

$$\mathrm{R}^{n_1+2t} \left\{ 1 + \frac{n_1+2t}{2}(1-\mathrm{R}^2) + \frac{(n_1+2t)(n_1+2t+2)}{2\,.\,4}(1-\mathrm{R}^2)^2 + \dots \right.$$
$$\left. + \frac{(n_1+2t)\dots(n_1+2t+n_2-4)}{2\,.\,4\dots(n_2-2)} \, (1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)} \right\}$$

or

$$\sum_{p=0}^{\frac{1}{2}(n_2-2)} \frac{(1-\mathrm{R}^2)^p}{p\,!} \cdot \frac{[\frac{1}{2}(n_1+2t+2p-2)]\,!}{[\frac{1}{2}(n_1+2t-2)]\,!} \, (\mathrm{R}^2)^{\frac{1}{2}(n_1+2t)}.$$

Again

$$\sum_{t=0}^{\infty} \frac{[\frac{1}{2}(n_1+n_2+2t-2)]!}{[\frac{1}{2}(n_1+n_2-2)]!\,t!}\,\rho^{2t}\,\frac{[\frac{1}{2}(n_1+2t+2p-2)]!}{[\frac{1}{2}(n_1+2t-2)]!}\,(\mathrm{R}^2)^{\frac{1}{2}(n_1+2t)}$$

is

$$\mathrm{R}^{n_1}\,\frac{[\frac{1}{2}(n_1+2p-2)]!}{[\frac{1}{2}(n_1-2)]!}\,\mathrm{F}\left(\frac{n_1+n_2}{2},\ \frac{n_1+2p}{2},\ \frac{n_1}{2},\ \rho^2\mathrm{R}^2\right)$$

or

$$\frac{[\frac{1}{2}(n_1+2p-2)]!}{[\frac{1}{2}(n_1-2)]!}\cdot\frac{\mathrm{R}^{n_1}}{(1-\rho^2\mathrm{R}^2)^{\frac{1}{2}(n_1+n_2+2p)}}\,\mathrm{F}\left(-p,\ -\frac{n_2}{2},\ \frac{n_1}{2},\ \rho^2\mathrm{R}^2\right),$$

which terminates in $p+1$ terms, and is equivalent to

$$\frac{\frac{1}{2}(n_2)!}{[\frac{1}{2}(n_2-2p)]!}\cdot\frac{\mathrm{R}^{n_1}(\rho^2\mathrm{R}^2)^p}{(1-\rho^2\mathrm{R}^2)^{\frac{1}{2}(n_1+n_2+2p)}}\,\mathrm{F}\left(-p,\ -\frac{n_1+2p-2}{2},\ \frac{n_1-2p+2}{2},\ \frac{1}{\rho^2\mathrm{R}^2}\right),$$

or

$$\frac{(\frac{1}{2}n_2)!}{[\frac{1}{2}(n_2-2p)]!}\cdot\frac{\mathrm{R}^{n_1}\,(-)^p}{(1-\rho^2\mathrm{R}^2)^{\frac{1}{2}(n_1+n_2)}}\,\mathrm{F}\left(-p,\ \frac{n_1+n_2}{2},\ \frac{n_2-2p+2}{2},\ \frac{1}{1-\rho^2\mathrm{R}^2}\right).$$

The probability integral, when n_2 is even, may therefore be written in the forms

$$(1-\rho^2)^{\frac{1}{2}(n_1+n_2)}\,\mathrm{R}^{n_1}\sum_{p=0}^{\frac{1}{2}(n_2-2)}\frac{[\frac{1}{2}(n_1+2p-2)]!}{[\frac{1}{2}(n_1-2)]!\,p!}\,\frac{(1-\mathrm{R}^2)^p}{(1-\rho^2\mathrm{R}^2)^{\frac{1}{2}(n_1+n_2+2p)}}\,\mathrm{F}\left(-p,\ -\frac{n_2}{2},\ \frac{n_1}{2},\ \rho^2\mathrm{R}^2\right),$$

or

$$\left(\frac{1-\rho^2}{1-\rho^2\mathrm{R}^2}\right)^{\frac{1}{2}(n_1+n_2)}\mathrm{R}^{n_1}\sum_{p=0}^{\frac{1}{2}(n_2-2)}(-)^p\,\frac{(\frac{1}{2}n_2)!}{[\frac{1}{2}(n_2-2p)]!\,p!}\,(1-\mathrm{R}^2)^p\,\mathrm{F}\left(-p,\ \frac{n_1+n_2}{2},\ \frac{n_1-2p+2}{2},\ \frac{1}{1-\rho^2\mathrm{R}^2}\right).$$

both of which terminate in $\frac{n_2}{8}(n_2+2)$ elementary terms.

When $n_2 = 2$, we have the simple probability integral

$$\{(1-\rho^2)/(1-\rho^2\mathrm{R}^2)\}^{\frac{1}{2}(n_1+2)}\,\mathrm{R}^{n_1};$$

when $n_2 = 4$, it becomes

$$\left(\frac{1-\rho^2}{1-\rho^2\mathrm{R}^2}\right)^{\frac{1}{2}(n_1+4)}\left\{\frac{n_1+4}{2}\,\frac{1-\mathrm{R}^2}{1-\rho^2\mathrm{R}^2}-(1-2\mathrm{R}^2)\right\}\mathrm{R}^{n_1},$$

and, when $n_2 = 6$,

$$\left(\frac{1-\rho^2}{1-\rho^2\mathrm{R}^2}\right)^{\frac{1}{2}(n_1+6)}\left\{\frac{(n_1+6)(n_1+8)}{2\,.\,4}\left(\frac{1-\mathrm{R}^2}{1-\rho^2\mathrm{R}^2}\right)^2 - \frac{n_1+6}{2}(2-3\mathrm{R}^2)\frac{1-\mathrm{R}^2}{1-\rho^2\mathrm{R}^2} + (1-3\mathrm{R}^2+3\mathrm{R}^4)\right\}\mathrm{R}^{n_1}.$$

It should be observed that the coefficient of $\{(1-\mathrm{R}^2)/(1-\rho^2\mathrm{R}^2)\}^p$ is given by

$$\frac{[\frac{1}{2}(n_1+n_2+2p-2)]\,!}{[\frac{1}{2}(n_1+n_2-2)]\,!\;p\,!^2}\,\frac{d^p}{dx^p}\left\{\frac{(\mathrm{R}^2+x)^{\frac{1}{2}n_2}-\mathrm{R}^{n_2}}{x}\right\}$$

when $x = -1$.

7. *Extension of the Analysis of Variance.*

The distribution of the simple correlation coefficient, although one of the first sampling distributions to be determined with exactitude[*], has hitherto occupied a somewhat isolated position. For all the exact distributions of statistics since discovered have grouped themselves in a single system; they are all amenable to the same technical procedure known as the analysis of variance; and all may be reduced to an equivalent problem of the distribution of the difference of the logarithms of two independent estimates of variance, based respectively upon n_1 and n_2 degrees of freedom.

The distribution of such an estimate s_1^2 derived from n_1 degrees of freedom is given by

$$df = \frac{1}{[\frac{1}{2}(n_1-2)]\,!}\,t_1^{\frac{1}{2}(n_1-2)}\,e^{-t_1}\,dt_1,\ \text{where}\ t_1 = \frac{n_1 s_1^2}{2\sigma^2},$$

and σ is the parameter of which s_1 is the first estimate.

If, now, $t_2 = n_2 s_2^2/2\sigma^2$, and

$$z = \log s_1 - \log s_2,$$

it follows that

$$t_1 = (n_1/n_2)\,e^{2z}\,t_2,$$

and the simultaneous distribution

$$df = \frac{1}{[\frac{1}{2}(n_1-2)]\,!}\,t_1^{\frac{1}{2}(n_1-2)}\,e^{-t_1}\,dt_1 \cdot \frac{1}{[\frac{1}{2}(n_2-2)]\,!}\,t_2^{\frac{1}{2}(n_2-2)}\,e^{-t_2}\,dt_2$$

may be written

$$df = \frac{2}{[\frac{1}{2}(n_1-2)]\,!\;[\frac{1}{2}(n_2-2)]\,!}\left(\frac{n_1}{n_2}e^{2z}\right)^{\frac{1}{2}n_1} t_2^{\frac{1}{2}(n_1+n_2-2)}\,e^{t_2\left(1+\frac{n_1}{n_2}e^{2z}\right)}\,dt_2\,dz\,;$$

---

[*] Fisher, 'Biometrika,' vol. 10, p. 507 (1915).

this expression may be integrated with respect to t_2 to yield the distribution of z, in the form

$$df = 2 \frac{[\frac{1}{2}(n_1+n_2-2)]!}{[\frac{1}{2}(n_1-2)]!\ [\frac{1}{2}(n_2-2)]!} \cdot \frac{n_2^{\frac{1}{2}n_2}\, n_1^{\frac{1}{2}n_1}\, e^{n_1 z}}{(n_2+n_1 e^{2z})^{\frac{1}{2}(n_1+n_2)}} \cdot dz,$$

completely independent of the unknown variance.

By the insertion of the appropriate values of n_1 and n_2, including the important bounding values of unity and infinity, the appropriate distribution of z for the analysis of variance is obtained. In the case, for example, of the multiple correlation coefficient drawn from uncorrelated material, n_1 is equated to the number of independent variates, $n_1 + n_2 + 1$ to the sample number, and $2z$ to

$$\log(\mathrm{R}^2/n_1) - \log(\overline{1-\mathrm{R}^2}/n_2).$$

It was from the first obvious that this system was capable without formal modification of extension to the case in which s_1 and s_2 were estimates of two different parameters σ_1 and σ_2, for in such cases we have only to write $\zeta = \log\sigma_1 - \log\sigma_2$, and the distribution found above will be that appropriate to the variate $z - \zeta$.

The new system of distributions found for the multiple correlation coefficient derived from correlated material is not only a generalisation of that previously found* for the simple correlation coefficient, but provides an extension of a different kind from that mentioned above to the analysis of variance. For the limiting distribution found in section 5 (distribution of B) may be interpreted as the distribution of the sum of the squares of n_1 variates normally distributed with equal variance, but not with zero means as in all cases previously discussed.

To show this, let $\mathrm{T} = \frac{1}{2\sigma^2}\overset{n_1}{\underset{p=1}{\mathrm{S}}}(x_p - a_p)^2$, in which $x_1, \ldots, x_{n_1}$ are variates distributed independently about zero with common variance σ^2. Let $\xi = \mathrm{S}(ax)/\sigma\mathrm{S}(a^2)$, then ξ will be normally distributed about zero with unit variance, and if we write $\frac{1}{2}\chi^2$ for T $\frac{1}{2}(\xi - \sqrt{\mathrm{S}(a^2)}/\sigma)^2$ or $\frac{1}{2}\overset{n_1}{\underset{1}{\mathrm{S}}}(x^2/\sigma^2) - \frac{1}{2}\xi^2$, which is the sum of the squares of $(n_1 - 1)$ quantities independently distributed about zero with unit variance, it appears that the distribution of χ^2 is of the familiar form

$$\frac{1}{[\frac{1}{2}(n_1-3)]!}(\tfrac{1}{2}\chi^2)^{\frac{1}{2}(n_1-3)}\, e^{-\frac{1}{2}\chi^2}\, d(\tfrac{1}{2}\chi^2),$$

and is independent of that of ξ, namely $(2\pi)^{-\frac{1}{2}}\, e^{-\frac{1}{2}\xi^2}\, d\xi$.

---

* Fisher, 'Biometrika,' vol. 10, p. 507 (1915).

If, now, β is written for $\sqrt{S(a^2)}/\sigma$ and x for $\xi - \beta$, it follows that

$$\tfrac{1}{2}\chi^2 = T - \tfrac{1}{2}x^2,$$

and as the same value of x^2 is provided by the two values of ξ, $\beta \pm \sqrt{x^2}$, the frequency element required from the distribution of ξ is

$$\frac{1}{\sqrt{2\pi}}\{e^{-\frac{1}{2}(\beta+x)^2} + e^{-\frac{1}{2}(\beta-x)^2}\}\,dx,$$

only positive values of x being now considered. Substituting for χ^2 in terms of T and x, the frequency distribution of the latter two variates will be given by

$$df = \frac{1}{[\frac{1}{2}(n_1-3)]\,!}(T-\tfrac{1}{2}x^2)^{\frac{1}{2}(n_1-3)}\,dT\,.\,\frac{2}{\sqrt{2\pi}}e^{-\frac{1}{2}\beta^2}e^{-T}\cosh(\beta x)\,.\,dx.$$

For a given value of T, the variate x cannot exceed $\sqrt{2T}$, and the random sampling distribution of T is therefore found by integrating between 0 and $\sqrt{2T}$. Expanding the hyperbolic cosine in powers of x, and integrating term by term, since

$$\int_0^{\sqrt{2T}}(T-\tfrac{1}{2}x^2)^{\frac{1}{2}(n_1-3)}\frac{x^{2p}\beta^{2p}}{(2p)\,!}\,dx$$
$$= \frac{[\frac{1}{2}(2p-1)]\,!\,[\frac{1}{2}(n_1-3)]\,!}{[\frac{1}{2}(n_1+2p-2)]\,!}T^{\frac{1}{2}(n_1+2p-2)}\,2^{\frac{1}{2}(2p-1)}\frac{\beta^{2p}}{(2p)\,!},$$

we have the distribution of T in the form.

$$df = e^{-\frac{1}{2}\beta^2}e^{-T}\sum_{p=0}^{\infty}\frac{T^{\frac{1}{2}(n_1+2p-2)}\beta^{2p}}{[\frac{1}{2}(n_1+2p-2)]\,!\,2^p\,.\,p\,!}\,dT,$$

or

$$df = e^{-\frac{1}{2}\beta^2}\frac{T^{\frac{1}{2}(n_1-2)}}{[\frac{1}{2}(n_1-2)]\,!}e^{-T}\left\{1+\frac{1}{n_1}(T\beta^2)+\frac{1}{n_1(n_1+2)}\frac{(T\beta^2)^2}{2\,!}+\,.\,.\right\}dT,$$

which is the B-distribution of section 5 if T is equated to $\tfrac{1}{2}B^2$.

This interpretation of the distribution previously obtained adds greatly to its importance, for it is seen to replace the χ^2 distribution of the analysis of variance for cases in which the sum of squares corresponding to n_1 degrees of freedom is derived theoretically for non-central deviations with fixed central displacements. This will be similar to, but not identical with, the case of the n_1 degrees of freedom in multiple correlation in its proper form; for although these are non-central, the displacements will depend on the variation *in the sample* of the independent variates, and this will vary from sample to sample. In many cases, however, such as the dependence of weather upon the position and altitudes of a number of fixed meteorological stations, we are not interested in the effects of possible variations in the positions of the stations, but solely in the possible variations of the weather at these spots. In fact, the problem of practical importance is often that in which the central displacements are constant, and although it may be urged, rightly enough, that to such cases the purely empirical concept of multiple correlation is not the most appropriate approach, yet it remains true that of the practical applications of multiple correlation methods many are of this kind.

The direct extension of the analysis of variance for non-central squares may be completed by writing

$$df = \frac{1}{[\frac{1}{2}(n_2-2)]\,!}\, t_2^{\frac{1}{2}(n_2-2)}\, e^{-t}\, dt_2 \quad \text{and} \quad \frac{\mathrm{T}}{t_2} = \frac{\mathrm{R}^2}{1-\mathrm{R}^2} = \frac{n_1}{n_2} e^{2z},$$

then, if, in spite of the caution above, we choose to express our results in terms of R,

$$t_2 = \frac{1-\mathrm{R}^2}{\mathrm{R}^2}\mathrm{T}, \quad dt_2 = -\frac{1-\mathrm{R}^2}{\mathrm{R}^2}\mathrm{T}\frac{d\,(\mathrm{R}^2)}{\mathrm{R}^2\,(1-\mathrm{R}^2)}, \quad t_2 + \mathrm{T} = \frac{\mathrm{T}}{\mathrm{R}^2},$$

and the distribution of R is found by integrating with respect to T from 0 to ∞, the expression

$$\frac{1}{[\frac{1}{2}(n_2-2)]\,!}\, e^{-\frac{1}{2}\beta^2}\, e^{-\mathrm{T}/\mathrm{R}^2} \left(\frac{1-\mathrm{R}^2}{\mathrm{R}^2}\right)^{\frac{1}{2}n_2} \frac{d\mathrm{R}^2}{\mathrm{R}^2\,(1-\mathrm{R}^2)} \sum_{p=0}^{\infty} \frac{\mathrm{T}^{\frac{1}{2}(n_1+n_2+2p-2)}\,\beta^{2p}}{[\frac{1}{2}(n_1+2p-2)]\,!\;2^p\,.\,p\,!}\,d\mathrm{T},$$

a process which yields

$$df = (\mathrm{R}^2)^{\frac{1}{2}(n_1-2)} \frac{(1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}}{[\frac{1}{2}(n_1-2)]\,!}\, e^{-\frac{1}{2}\beta^2} \sum_{p=0}^{\infty} \frac{[\frac{1}{2}(n_1+n_2+2p-2)]\,!}{[\frac{1}{2}(n_1+2p-2)]\,!} \frac{(\mathrm{R}^2\beta^2)^p}{2^p\,.\,p\,!}\, d\,(\mathrm{R}^2),$$

or

$$df = \frac{[\frac{1}{2}(n_1+n_2-2)]\,!}{[\frac{1}{2}(n_1-2)]\,!\,[\frac{1}{2}(n_2-2)]\,!} (\mathrm{R}^2)^{\frac{1}{2}(n_1-2)}\, (1-\mathrm{R}^2)^{\frac{1}{2}(n_2-2)}\, e^{-\frac{1}{2}\beta^2} \left\{1 + \frac{n_1+n_2}{n_1\,.\,1\,!}\frac{\mathrm{R}^2\beta^2}{2} + \frac{(n_1+n_2)\,(n_1+n_2+2)}{n_1\,(n_1+2)\,.\,n\,2\,!}\left(\frac{\mathrm{R}^2\beta^2}{2}\right)^2 + \ldots\right\} d\,(\mathrm{R}^2), \quad \text{(C)}$$

a third general distribution of this interesting group.

Although it will not be possible within the limits of this paper to give an account of the properties of the distribution of Type (C), beyond indicating their analogy with those of Type (A), it should not be overlooked that in the problems in which the multiple correlation coefficient is actually employed, distributions of Type (C) will be, owing to the absence or irrelevance of sampling variation in the variances of the independent variates, of at least as frequent occurrence as those of Type (A).

A typical example of the distinction here drawn is provided by the correlation ratio. If corresponding to any value x of the independent variate a number of values n_x of the dependent variate y is observed, then the correlation ratio η^2 of y on x is defined by the relation

$$\frac{\eta^2}{1-\eta^2} = \frac{S\{n_x(\bar{y}_x - \bar{y})^2\}}{S(y - \bar{y}_x)^2},$$

in which $\bar{y}_x$ is the mean of y in any array, and $\bar{y}$ is the general mean; the variance in all arrays is supposed equal, and the summation in the numerator is applied to the several arrays, while that in the denominator is applied to the whole of the individual observations. In most practical cases the idea of a sampling distribution of η^2 can only be given a definite meaning by supposing the number n_x in each array to be the same for all samples. In such a case the distribution of η^2 will be that of R^2 in distribution (C), with n_1 equal to one less than the number of arrays, and $n_1 + n_2 + 1$ equal to the total number of observations. If, however, the numbers n_x be regarded as subject to sampling variations, then the distribution (A) may be used, and will be exact, apart from grouping errors, if the expectations of y for the values of x in the sampled population are normally distributed.

Summary.

By an appropriate linear transformation of the independent variates it may be shown that the sampling distribution of the multiple correlation coefficient does not depend on the whole matrix of correlations between these variates, but solely upon the multiple correlation in the population sampled

The actual distribution (A) may then easily be obtained by similar methods to those by which the distribution of the simple correlation coefficient has been obtained.

The frequency function involves a hypergeometric function of $\rho^2 R^2$ which is a rational function when n_1 and n_2 are both even, algebraic when n_2 only is even, and reducible to circular functions when n_1 and n_2 are both odd.

The case of large samples yields a series of distributions (B) of great interest, involving Bessel functions, which connect the χ^2 distributions with the Gaussian, and are intimately related to a double Poisson summation. Owing to the practical importance of this limiting form a table of its 5 per cent. points is given up to seven independent variates.

When n_2 is even, the probability integral of the general distribution is expressible in finite terms which are developed in section 6.

The (B) distribution of Section 5 replaces the χ^2 distribution in the analysis of variance if the squares summed are non-central. An analysis of variance so extended leads to a third group of distributions (C), closely related to (A), and tending like it to a common limit (B). The distinction between (A) and (C) arises from the fact that in cases proper to the multiple correlation the central displacements will vary from sample to sample owing to variations in the second order moment coefficients of the independent variates, and for such cases (A) is the correct distribution. The type (C), however, is of frequent occurrence owing to the absence or irrelevance of such variation.

The Scattering Power of a Bare Nucleus according to Wave Mechanics.

By G. Temple, Ph.D., Imperial College, 1851 Exhibition Research Student.

(Communicated by S. Chapman, F.R.S.—Received September 22, 1928.)

Mott's investigation* of the scattering of an infinite plane wave by a bare nucleus involves rather delicate considerations of the behaviour of a power series near its circle of convergence, and it therefore appears desirable to obtain his result by a more direct and simple method. Such a method is given in this paper.

The scattering power of a nucleus may be defined as the volume density of electricity in the scattered wave when the volume density of the incident wave is unity around the nucleus.

§ 1. *Elementary Solutions of the Wave Equation in Paraboloidal Co-ordinates.*

The analysis is made simpler by replacing the spherical polar co-ordinates (r, θ) by the paraboloidal co-ordinates (ξ, η) where

$$\xi = r(1 + \cos\theta), \qquad \eta = r(1 - \cos\theta).$$

In these co-ordinates the wave equation for waves with symmetry about the z-axis is

$$\frac{\partial}{\partial\xi}\left\{\xi\frac{\partial\psi}{\partial\xi}\right\} + \frac{\partial}{\partial\eta}\left\{\eta\frac{\partial\psi}{\partial\eta}\right\} + \tfrac{1}{4}\kappa^2(\xi + \eta)\psi + \gamma\psi = 0, \tag{1.1}$$

---

* 'Roy. Soc. Proc.,' A, vol 118, p. 542 (1928)

where $2\pi/\kappa$ is the wave-length of the incident beam and γ is the reciprocal of the radius of Bohr's "ground-orbit" for a nucleus of charge Ze, *i.e.*,

$$h\kappa = 2\pi m v \quad \text{and} \quad h^2\gamma = 4\pi^2 m Z e^2, \tag{1.2}$$

v being the velocity of projection of the associated stream of electrons.

The elementary solutions of the wave-equation are of the form

$$\exp(\iota\kappa z)\,\mathrm{L}(-i\kappa\xi\,;\ \nu + i\alpha)\,\mathrm{L}(i\kappa\eta\,;\ \nu - i\alpha), \tag{1.3}$$

where ν is any complex constant, $\alpha = \gamma/2\kappa$, and $\mathrm{L}(u\,;\ n)$ is a solution of the equation,

$$u\frac{d^2\mathrm{L}}{du^2} + (1-u)\frac{d\mathrm{L}}{du} + n\mathrm{L} = 0. \tag{1.4}$$

A fundamental system of solutions of this equation is furnished by the functions

$${}_1\mathrm{F}_1(-n\,;\ 1\,;\ u) = \frac{1}{\Gamma(-n)}\sum_{m=0}^{\infty}\frac{\Gamma(m-n)\,.\,u^m}{\Gamma^2(m+1)} = \mathrm{F}(u\,;\ n),$$

and

$$(-)^{-n-1}(-u)^{-\frac{1}{2}}e^{\frac{1}{2}u}\mathrm{W}_{-n-\frac{1}{2},0}(-u) = \frac{e^u}{\Gamma(n+1)}\int_0^{\infty} e^{-t}t^n(u-t)^{-n-1}\,dt$$
$$= \mathrm{G}(u\,;\ n).$$

Provided that $\mathrm{R}(n+1) > 0$, the integral defining $\mathrm{G}(u\,;\ n)$ converges uniformly in any region excluding the origin, $u = 0$, where there is a logarithmic singularity. The asymptotic expansion of this function is

$$u^{-n-1}e^u\,{}_2\mathrm{F}_0(n+1,\ n+1\,;\ u^{-1}).$$

The series defining $\mathrm{F}(u\,;\ n)$ converges uniformly for all finite values of u and its asymptotic expansion* is

$$\frac{(-u)^n}{\Gamma(n+1)}\,{}_2\mathrm{F}_0(-n,\ -n\,;\ -u^{-1}) + \frac{e^u u^{-n-1}}{\Gamma(-n)}\,{}_2\mathrm{F}_0(n+1,\ n+1; u^{-1}).$$

The main point of the analysis is that both parts of this expansion must be retained when u is a pure imaginary. We note that the solution of equation (1.4) whose asymptotic expansion is

$$(-u)^n\,{}_2\mathrm{F}_0(-n,\ -n\,;\ -u^{-1})/\Gamma(n+1),$$

is

$$\mathrm{F}(u\,;\ n) - \mathrm{G}(u\,;\ n)/\Gamma(-n) = \mathrm{H}(u\,;\ n), \text{ say.}$$

§ 2. *The Scattering of a Plane Wave.*

We proceed to obtain the only elementary solution (1.3) which is physically admissible.

---

* Barnes, 'Trans. Camb. Phil. Soc.,' vol. 20, p. 253 (1906).

In the first place we note that we must exclude terms involving $\exp(-i\kappa r)$ (which correspond to a converging wave) from the asymptotic expansion of the solution. Hence we are restricted to solutions of the form

$$\exp(i\kappa z)\, \mathrm{H}(-i\kappa\xi,\ \nu + i\alpha)\, \mathrm{F}(i\kappa\eta;\ \nu - i\alpha).$$

In the second place we note that the solution will present a singularity along the negative half of the z-axis ($\xi = 0$) unless $\nu = -i\alpha$. Our solution thus reduces to

$$\exp(i\kappa z)\, \mathrm{F}(i\kappa\eta;\ -2i\alpha). \tag{2.1}$$

Along the positive half of the z-axis ($\eta = 0$) this function becomes simply $\exp(i\kappa z)$, but provided $\theta \neq 0$ it follows from the results of § 1 that its asymptotic expansion is

$$\frac{\exp\{i\kappa z - 2i\alpha \log \kappa (r - z) - \pi\alpha\}}{\Gamma(1 - 2i\alpha)}\, {}_2\mathrm{F}_0(2i\alpha,\ 2i\alpha,\ i/\kappa\eta)$$
$$+ \frac{\exp\{i\kappa r + 2i\alpha \log \kappa (r - z) - \pi\alpha\}}{i\kappa\eta\, \Gamma(2i\alpha)}\, {}_2\mathrm{F}_0(1 - 2i\alpha,\ 1 - 2i\alpha;\ -i/\kappa\eta). \tag{2.2}$$

The first term represents the incident wave, the second term the wave scattered by the nucleus.

The ratio of the moduli of the leading terms in the asymptotic expansions of the scattered and reflected waves is

$$\left|\frac{\Gamma(1 - 2i\alpha)}{i\kappa\eta\, \Gamma(2i\alpha)}\right| = \frac{2\alpha}{\kappa\eta} \left|\frac{\Gamma(1 - 2i\alpha)}{\Gamma(1 + 2i\alpha)}\right| = \gamma/2\kappa r \sin^2 \tfrac{1}{2}\theta.$$

Introducing the values of γ and κ (1.2) we find that the scattering power of the bare nucleus is

$$(Ze^2/2mv^2r)^2 \operatorname{cosec}^4 \tfrac{1}{2}\theta \qquad (\theta \neq 0). \tag{2.3}$$

This is the classical value obtained by Rutherford.

The New Metric of Einstein and the Wave Equation.

H. T. FLINT, Reader in Physics, University of London, King's College.

(Communicated by O W Richardson, F.R S -Received October 10, 1928)

In two papers published in these ' Proceedings '* the suggestion was made that the wave equation is a metrical equation and that the quantity ψ is characteristic of the metrics of physical space.

In developing this view a definition of equality of vectors at different points was given and ψ was interpreted as the natural scale factor Except for the definition of equality the space was Riemannian. The system was suggested by that of Weyl and Eddington and was, in fact, almost identical with it.

The wave equation turned out to be very simple in form and a number of results were derived which seemed to show that in the search for the interpretation of ψ a step had been taken in the right direction. There is, however, the difficulty that it does not appear possible to go far enough in other directions with this metrical scheme. It may be that the scheme is not a physical one and that the natural geometry and metrics of space is not that proposed.

On applying the same method to a five-dimensional continuum the wave equation is even simpler than in four and we have other advantages. The tracks of particles in this space are geodesics, whether they are under gravitational or electromagnetic forces, and it is merely an artificial device to separate gravitational and electromagnetic phenomena. Thus the relativistic ideal is attained and the wave equation finds a metrical interpretation.

It has been shown that by introducing the quantity ψ the laws of gravitation and electromagnetism together with the wave equation may be united in a single variational principle. It may then be possible to derive a single geometrical equation which will unite all these laws just as the law of gravitation finds a geometrical interpretation in the theory of Relativity. Finally, processes which are continuous in five dimensions may appear to be discontinuous in four, so that if we regard the physicist as a four-dimensional observer, experimenting in time and space, we may regard his discontinuities as due to his inability to perceive more than four dimensions.

From the point of view of mathematical form the five-dimensional view has a strong claim, but an objection to the introduction of a fifth dimension is natural, especially as it may never be possible to observe it. We may turn again to a four-dimensional continuum and enquire if our difficulties can be removed by the adoption of a new system of geometry.

---

* Vol. 117, pp. 625, 630 (1928).

The new metric of Einstein* is proposed with this end in view and the method he adopts is to add to the Riemannian geometrical scheme a definition of parallelism at separate points.

A system of orthogonal unit vectors is set up at each point of space and at first there is no relation between the orthogonal system at one point and that at another. Let one point be P and another Q, then if the components of a vector A at P along the orthogonal system at P are equal to the components of a vector B at Q along the orthogonal system at Q, the vectors are parallel These vectors are also equal in magnitude, so that the definition makes it possible to equate vectors at the two points. There is nothing to fix the orientation of the orthogonal system at Q with respect to that at P, and if physical phenomena are to be described in this system they must be expressed by equations which are unaffected by rotation of the axes. This is the new principle of invariance and gives rise to new vectors and tensors possessing this property. Each orthogonal component, ε_a, say, has components along the natural axes appropriate to the space, which are in general curved lines. These components are denoted by $(h^1{}_a, h^2{}_a, h^3{}_a, h^4{}_a)$ where a has the values 1 to 4, so that there are 16 of these quantities. Associated with them are the covariant components $(h_{1a}, h_{2a}, h_{3a}, h_{4a})$. These quantities are fundamental quantities of the metric and are to be compared to the components g^{mn} and g_{mn} of the Riemannian theory, which now turn out to be functions of the h's, the relations being

$$\left.\begin{aligned} g_{mn} &= h_{ma}h_{na} \\ g^{mn} &= h^m{}_a h^n{}_a \end{aligned}\right\}, \tag{1}$$

the summation being taken over the values of a.

In addition to these we have the relations :

$$\left.\begin{aligned} h_{ma}h^n{}_a &= \delta^n{}_m \\ h^m{}_a h_{mb} &= \delta_{ab} \end{aligned}\right\}. \tag{2}$$

It is our object here to consider the place of the wave equation in this metric and we can do this in the same way as in the former papers.

The wave equation in its general form, for an electron in a gravitational and electromagnetic field, is ·

$$\Box\psi - \frac{4\pi ie}{k}\phi^a\frac{\partial\psi}{\partial x^a} - \frac{4\pi^2}{h^2}(e^2\phi_a\phi^a - p_a p^a) = 0. \tag{3}$$

---

* 'Sitzb. Preuss. Akad. Wiss.,' vol. 17, p. 217 (1928).

The view which we have taken is that this is of the form:

$$\operatorname{div}(\operatorname{grad}\psi) = A\psi, \tag{4}$$

where A is zero or equal to a quantity composed simply of the generalised momentum Π_m, where Π_m is usually interpreted as $(p_m + e\phi_m)$

In the five-dimensional metric A is zero and the wave equation takes a particularly simple form.

In the metric of Einstein we shall consider the operation of divergence and apply it to ψ as in equation (4). We shall see that in this case the equation, div (grad ψ) $= 0$, gives us exactly the wave equation (3), throws light on the physical interpretation of fundamental quantities of the metric and reveals ψ as a metrical quantity.

Einstein shows that a vector A at a point P, of which A^m is a typical component, will be equal to a vector B at Q if

$$B^m = A^m + dA^m,$$

where

$$dA^m = -\Delta^m_{nk} A^n \, dx^k. \tag{5}$$

and where

$$\Delta^m_{nk} = -h^m{}_a \frac{\partial h_{na}}{\partial x^k},$$

the h's having the interpretation given above

Δ^m_{nk} plays a part corresponding to the Riemannian quantity Γ^m_{nk}.

The equation (5) corresponds to the Riemannian parallel displacement, which we write:

$$\bar{d}A^m = -\Gamma^m_{nk} A^n \, dx^k. \tag{6}$$

An important difference is that a parallel displacement in the Einstein metric when taken round a closed curve leads back to the original vector, while this is not so in the Riemannian metric.

The difference between the vectors in (5) and (6) is:

$$dA^m - \bar{d}A^m = (\Gamma^m_{nk} - \Delta^m_{nk}) A^n \, dx^k, \tag{7}$$

and to the tensor $(\Gamma^m_{nk} - \Delta^m_{nk})$, Einstein attaches a fundamental character.

In his paper on the physical significance of the theory he divides this tensor into its symmetric and antisymmetric parts, viz.,

$$\Gamma^m_{nk} - \tfrac{1}{2}(\Delta^m_{nk} + \Delta^m_{kn}) \quad \text{and} \quad \tfrac{1}{2}(\Delta^m_{kn} - \Delta^m_{nk}).$$

The second of these is denoted by Λ^m_{kn} and the electromagnetic potential

ϕ_k is identified with $\Lambda^m{}_{km}$. Maxwell's equations can then be derived from a variational theorem.

It follows from the definition of equality, (5), that the variation of a vector component in this space is given by:

$$\frac{\partial A^m}{\partial x^k} + \Delta^m{}_{nk} A^n,$$

just as in Riemannian space it is

$$\frac{\partial A^m}{\partial x^k} + \Gamma^m{}_{nk} A^n,$$

hence the expression for divergence is

$$\frac{\partial A^m}{\partial x^m} + \Delta^m{}_{nm} A^n.$$

With reference to covariant components we have:

$$\frac{\partial}{\partial x^m}(g^{ms}A_s) + \Delta^m{}_{nm} g^{ms} A_s = \left(g^{ms}\frac{\partial A_s}{\partial x^m} - g^{ms}\Gamma^t{}_{ms}A_t\right) - g^{ms}(\Gamma^r{}_{sr} - \Delta^r{}_{sr})\, A_m,$$

where the quantity in the first pair of brackets on the right is the Riemannian divergence and if A_s is written as a gradient it becomes the generalised Laplacian expression $\Box^2 \psi$, where $A_s = \dfrac{\partial \psi}{\partial x^s}$.

If then we consider the new expression for div (grad ψ) we have

$$\text{div}\,(\text{grad}\ \psi) = \Box^2 \psi - g^{ms}(\Gamma^r{}_{sr} - \Delta^r{}_{sr})\frac{\partial \psi}{\partial x^m}. \tag{8}$$

Let us now write $\dfrac{1}{\psi}\dfrac{\partial \psi}{\partial x^m} = a\Pi_m$ and equate (8) to zero.

Then in order to compare the resulting equation with (3) we will write it in the form:

$$\Box^2 \psi - \frac{4\pi i e}{h}\phi^m \frac{\partial \psi}{\partial x^m} - \left\{g^{ms}(\Gamma^r{}_{sr} - \Delta^r{}_{sr}) - \frac{4\pi i e}{h}\phi^m\right\}\frac{\partial \psi}{\partial x^m} = 0. \tag{9}$$

In order that this may be identical with (3) we must equate $(\Gamma^r{}_{sr} - \Delta^r{}_{sr})$ to the generalised momentum multiplied by a constant a which is equal to $2\pi i/h$, for then the third term of the equation becomes:

$$a^2\Pi_m(\Pi^m - 2e\phi^m) = \frac{4\pi^2}{h^2}(e^2\phi_m\phi^m - p_m p^m),$$

and (9) is identical with (3).

In addition to throwing the wave equation into the simple form:

$$\operatorname{div}(\operatorname{grad}\psi) = 0,$$

we learn that physically $(\Gamma^r_{sr} - \Delta^r_{sr})$ is to be interpreted as the generalised momentum and that actually

$$\Gamma^r_{sr} - \Delta^r_{sr} = a\Pi_s. \tag{11}$$

This appears to be consistent with Einstein's introduction of ϕ_k by means of Λ^m_{km}; for the expression on the left may be written thus

$$\Gamma^r_{sr} - \tfrac{1}{2}(\Delta^r_{sr} + \Delta^r_{rs}) + \tfrac{1}{2}(\Delta^r_{rs} - \Delta^r_{sr}),$$

and the last of these is $-\Lambda^r_{sr}$ or $-\phi_s$ according to Einstein. Equation (11) suggests that Λ^r_{sr} should be identified with $ae\phi_s$ and the introduction of the constant factor of ϕ_s does not appear to cause any difficulty in the derivation of Maxwell's equations The first two terms must then be regarded as representing the momentum multiplied by $2\pi i/h$, but this artificial division of Π_s into two parts is perhaps unnecessary.

Expressed in terms of the components, h, the left side of (11) is equal to

$$h^r_a\left(\frac{\partial h_{ra}}{\partial x^s} - \frac{\partial h_{sa}}{\partial x^r}\right),$$

which follows from (1) and (2) and from the value of Δ^m_{nk}, which is

$$h^m_a\frac{\partial h_{na}}{\partial x^k}.$$

Thus from (11) we have

$$h^r_a\left(\frac{\partial h_{ra}}{\partial x^s} - \frac{\partial h_{sa}}{\partial x^r}\right) = \frac{a}{\psi}\frac{\partial\psi}{\partial x^s} = a\frac{\partial}{\partial x^s}(\log\psi), \tag{12}$$

which is a relation that the components, h, must satisfy.

In the metric Einstein imposes no condition upon the components but it would appear that in a physical continuum it should be possible to associate the orthogonal systems at neighbouring points and this would imply some relation between the components.

In his paper introducing the metric he considers a case where the components, h_{ma}, are derived from a scalar quantity and it appears from (12) that the relation between the components is determined by ψ. The quantity on the left of (12) is a covariant vector formed from the product of the vector h^r_a, and the curl of h_{ra}, this product is the gradient of log ψ multiplied by $2\pi i/h$.

Summary.

The wave equation in the new metric of Einstein is shown to have a very simple form. The physical interpretation of a fundamental vector occurring in the system is deduced from the equation and the metrical character of the wave function and its equation is discussed.

OBITUARY NOTICES.

CONTENTS

PAGE

CHRISTIAN FELIX KLEIN (with portrait) i

HENDRIK ANTOON LORENTZ (with portrait) xx

THEODORE WILLIAM RICHARDS (with portrait) xxix

SIR JOHN ISAAC THORNYCROFT (with portrait) xxxv

F. Klein

CHRISTIAN FELIX KLEIN—1849-1925

At Düsseldorf, on the night of April 25th, 1849 there was anxiety in the house of the secretary to the Regierungspräsident Without, the cannon thundered on the barricades raised by the insurgent Rhinelanders against their hated Prussian rulers Within although all had been prepared for flight there was no thought of departure , on that night was born a son to the stern Prussian secretary That son was Felix Klein His birth was marked by the final crushing of the revolution of 1848 his life measured the domination of Prussia over Germany, and typifies all that was best and noblest in that domination , with his last illness came the consummation of its downfall Gradually, but irresistably, the nervous malady of which from time to time he had had serious warnings, mastered and prostrated him It was, he thought, partly due to heredity on his mother's side partly to his own unbridled expenditure of mental energy—an energy which remained so indomitable that, even during the last two years of his life, when he lay helpless, becoming daily weaker and weaker in body he never complained and remained clear to the end working and even correcting proof sheets At half past eight on the evening of Monday June 22nd, 1925, he passed painlessly away

2 Few mathematicians have left such ample material for forming an opinion of their life and work as Felix Klein We have his life, written by his own hand two years before his death * We have his Collected Mathematical Papers, in three volumes,† thoroughly revised by himself, and interspersed with supplementary notes and introductory articles of an autobiographical character We already have the greater part of his mathematical lectures in print, lectures which had for many years enjoyed a considerable publicity in lithographed form we have even a faithful record of lectures given by him in the years just preceding his death, carefully annotated by his colleague and successor, Professor Courant

After his death there appeared one after another, a number of sketches of the man and his work from the pens of many of his pupils But, just as a photograph of a man of unusual personality, or of a place of striking beauty, conveys little to one not personally acquainted with the original, so it is, and

---

* 'Göttinger Professoren Lebensbilder von eigener Hand No 4, Mitteilungen des Universitätsbundes Göttingen.' Jahrg 5, Heft 1 This will be referred to as "Autobiography"

† Referred to below as 'Ges Math. Abh'

so it must be, with these sketches of Klein. At the same time it behoves the Royal Society to attempt more than merely to chronicle the leading events in the life of him who was its oldest Foreign Member. It is not merely that for 40 years Klein figures in the List of the Society, and received in 1912 the Copley Medal. It is still less that he was personally known to and valued by many of the Fellows. The life of Klein is bound up with the regeneration of mathematical study.

3. We have been told* that Klein was one of the greatest mathematicians of all time, a claim which he would certainly not have countenanced; but his personal influence was as great, or greater, even outside his own country, than that, perhaps, of any mathematician of modern times. He owed this to his forceful and attractive personality, to his wide mathematical outlook and to his objective open-mindedness, only occasionally tinged by the play of personal feeling, for Klein had by no means the cold, calculating nature, supposed typical of the mathematician. He owed it also partly to that intangible agent, so often and so callously cited in England, "Luck," which afforded him, by a succession of unforeseen external events, the opportunity without which a man of the highest intellectual and moral endowments may remain mute inglorious. He owed it partly to his success in interesting his audiences, even in branches of mathematics of which his own command was comparatively slight. And he owed it to his untiring devotion to the cause of education; but his efforts in this direction were rendered effective by the instrumentality of one in whom Klein's ideas aroused a keen interest, Althoff, who from the Curatorship of the University of Göttingen, was transferred to a leading post in the Prussian Ministry.

4. Klein's genius had been precocious. At 17 he was chosen by Julius Plücker as his assistant in his physical laboratory at Bonn; that laboratory where Sommerfeld tells us, Plücker had invented, and Geissler, his glassblower, had fabricated, what in England is known as the Crooke's tube. All through his life Klein was a would-be physicist; like Goethe's Mephistopheles, who aimed at evil and complained that he always did good, Klein was all his life striving to do anything rather than Pure Mathematics, and found himself, with the shortest of interludes, doing nothing else. Indeed, Plücker's own mind had reverted, after an interval of some years, to his early geometrical interests. In his youth he had discovered the famous equations bearing his name, in his later years he became the inspirer of Klein.

It is fitting for the Royal Society to recall that it was Plücker, and not Cayley, who introduced the six homogeneous line-co-ordinates in space, connected by an identity of the second degree, in a paper in the *Transactions* of this Society (1865, pp. 725–791). All Klein's early work was in line geometry, or was related with that subject. Yet even in writing his Doctor's Dissertation, which he evolved at the age of 19, a year after Plücker's death, he unknowingly introduced, what was so characteristic a feature of his later work, the idea of a Group,* and the same is true of his treatment of one of the five theses which he had to uphold at his examination for the Doctor's degree; indeed, in order to show† that the quadratic complex which Battaglini had previously considered was not the most general quadratic complex, he used an argument based on the existence of transformations of Plücker's co-ordinates, involving three parameters, by means of which transformations the whole set of Battaglini's complexes remained unaltered; these transformations form, of course, a group.

---

* F. Pfeiffer.

It was in consequence of Plücker's death (22nd May, 1868), leaving unfinished his "New Geometry of Space, founded on the straight line as element," that the task of completing and issuing the second half of the book was entrusted by Clebsch to Plücker's young assistant, the only person, in fact, who, from constant intercourse with the master, was in a position to edit the manuscript and fill in the lacunæ in the spirit of the author.

5. Clebsch had recently been installed at Gottingen; thither Klein followed, and was thus brought into daily intercourse with him.‡ Under the influence of Clebsch, Klein developed his love for geometrical models, already excited by Plücker, who told him, by the way, that it was chiefly through Faraday that he had come to appreciate the importance of models. Faraday, as a non-mathematician, had found the construction of models invaluable in enabling him to comprehend that part of the mathematics of his day which he needed in his work. In this connexion we remark that Plücker, by his intimacy with English scientists, seems to have been instrumental in leading Klein to study the English language and the writings of English mathematicians, and so led to that friendship with the English-speaking nations and individuals to which Klein was loyal to the end.

It was at a meeting of mathematicians in the Bergstrasse at Göttingen in 1868 that Klein first saw an actual model (Wiener's) of the cubic surface with the 27 real straight lines the existence of which had been discovered in 1849 by Cayley, with whose work Klein was already familiar.§ This incited Klein to pursue the subject of models further himself, in which resolution he had Clebsch's approval, and he actually constructed (1871) models in zinc of the Kummer surface of the fourth degree with sixteen double points and sixteen double planes, as well as other particular cases of the same surface, which occur in line geometry. Throughout his life his interest in the subject of the real forms of curves and surfaces remained lively, and he believed intensely in the utility of models and actual constructions for didactic purposes. He was proud to be able to say at the end of his life that, thanks to his efforts and those of his friends, the Brills and Dyck, among others, no German university was without a proper collection of mathematical models.*

---

* 'Ges. Math. Abh.,' vol. 1, p. 18.

† *Ibid.*, p. 3.

‡ *Ibid.*, p. 51.

§ *Ibid.*, vol. 2, p. 3.

6. Klein's first stay at Göttingen was from the beginning of 1869 to the middle of August of the same year. His second stay was from the beginning of 1871 to the middle of September, 1872, and his final migration thither was in 1886. Between the first and the second stay there fall three important periods, his stay in Berlin, his stay at Paris, to which we shall return later, and the Franco-German war, in which he only took part in the ambulance corps, being refused for active service, and from which he was sent home invalided with typhoid fever.

In August, 1869, against the advice of Clebsch, which agreed with the caution Plücker had formerly given, Klein went to Berlin and stayed there till March, 1870. The auguries of Clebsch and Plücker were so far verified that Klein did not succeed in impressing his personality on the Berlin authorities, as he had done at Bonn and at Göttingen. Indeed, in spite of his work on Kummer's surface, and although he spoke on line geometry several times at Kummer's Seminar, he says himself,† and expresses surprise at the fact, that no link established itself between himself and Kummer's researches into algebraic systems of rays. In the mind of Weierstrass, a veritable antagonism to Klein seems to have arisen. The word went round among Weierstrass's pupils that Klein was "anathema maranatha," and that what he did was not mathematics at all. Yet, at the suggestion of Lipschitz, to whom he had submitted his Dissertation, Klein had grafted on the original sketch a discussion of degenerate cases, based on Weierstrass's Elementarteile; and, with his ardent desire for breadth of outlook and his determination not to be confined to any one school of thought, it must have been largely in hopes of profit from personal contact with Weierstrass that he insisted on going to Berlin. Yet he himself tells us† that he hardly attended any lectures, and that it was only through a fellow student, Kiepert, that he came across, and studied, Weierstrassian Elliptic Functions. He spoke once, and only once, at Weierstrass's Seminar; his subject, was Cayley's Theory of the Absolute, and, in conclusion, he threw out the suggestion that there was here a connection with Non-Euclidian Geometry, he having just heard of Non-Euclidian Geometry for the first time from another fellow student, Otto Stolz. Weierstrass absolutely rejected this suggestion, laying it down that the concept of distance was at the basis of all geometry. Later, Klein developed his idea in a series of papers extending over several years; he was able also, on the occasion of his attending the Meeting of the British Association at Bradford, in 1873, to discuss the question face to face with Cayley himself, as well as with Clifford and Sir Robert Ball. He never succeeded in convincing Cayley and Ball that he was not arguing in a circle,* and, indeed, to make the argument really convincing required a mind more patient and diffident than that of the impetuous and daringly intuitive Klein.

---

* 'Ges. Math. Abh.,' vol. 1, p. 51.

† Autobiography, p. 15.

7. It was the influence of Clebsh that obtained for Klein, at the early age of 23, the full Chair of Mathematics at the University of Erlangen. This would, however, not have afforded him sufficient opportunity of distinguishing himself, but for what seemed at first a disaster for the rising young geometer, the premature death of his chief supporter Clebsch from diphtheria on November 7th, 1872. Two days previously, Klein had given his first lecture to an audience of two, only one of whom reappeared, and he only once or twice. But the mantle of Clebsch could not fall otherwise than on the shoulders of Klein, and a number of advanced students, some older than Klein himself, migrated at once from Göttingen to Erlangen. The young professor found himself suddenly the centre of an intense mathematical life, such as had previously existed at Göttingen.†

It was at Erlangen that Klein produced what he himself regarded as his most notable achievement, the so-called Erlangen Programme. It would be no exaggeration to say that it has revolutionised the treatment of Geometry. Only in details has exception been taken to any part of it, which is the more remarkable when we consider the age of the author and the state of mathematics at the time when it was composed. The notes which Klein himself has added, and the account which he has given‡ of its production, have greatly increased its interest. The new idea which lies at its basis was that all the various species of Geometries which, during the 19th century in particular, had multiplied exceedingly (metrical geometry, projective geometry, line geometry, etc.) could be regarded from a single standpoint, that of the Theory of Groups, each different geometry being conceived as the theory of the invariants of an appropriate group. This idea, which as then formulated, seemed, even to Sophus Lie, "very astonishing,"§, was an abiding one in the mind of Klein, who thought he foresaw its extension to regions of mathematics other than geometry. In particular, the development by Einstein of the Theory of Relativity revived in his later years Klein's interest in his Erlangen Programme, and it was with some mortification that he found among his physicist friends no interest whatever in the possibility of a physical interpretation of the most general group of conformal transformations of four dimensional space, to which attention had been called by Bateman and Cunningham as leaving invariant the Maxwell-Lorentz equations.*

---

* 'Ges. Math. Abh.,' vol. 1, p. 242.

† Autobiography, p. 18.

‡ 'Ges. Math. Abh.,' vol. 1, pp. 411—414.

§ Sophus Lie, by F. Engel. 'Jahresber. d. deutschen Math. Ver.,' vol. 8, p. 39.

It was at Erlangen that Klein married.† His wife's beauty was conspicuous even in later life; her sweet and cultivated personality as Hausfrau, and the sad affliction of her growing deafness, which, when combined with the abnormally acute hearing of her husband, formed a sensible factor in the home-life, will always be remembered by those who had the privilege of enjoying the simple, old-fashioned hospitality of Klein's house. She was Anna Hegel, daughter of an Erlangen professor and grand-daughter of the philosopher Hegel.

8. At Erlangen, Klein worked on the lines which he had drawn up in his Programme, taking as guiding star a simple principle :— "Algebraic forms, which are transformed into themselves by finite groups of linear substitutions, have special properties which are in consequence easily perceived." This he illustrated, in the first instance, by the binary forms of the third and fourth degree, using Riemann's interpretation of the complex variable $x + iy$ on the sphere. We feel here at once the influence which Clebsch had exerted on the mind of Klein, in directing his attention to linear transformations and the theory of invariants and covariants of binary forms. But Klein had seized on this subject-matter with characteristic breadth of view, and he tells us‡ that he had all along asked himself, with a mixture of hope and fear, what Clebsch would say to a presentation of the subject of geometry so essentially different from his own systematically projective method; but Clebsch never saw the Erlangen Programme.‡

In the peroration of the Programme, Klein, who had come to some knowledge of the work of Galois during his stay in Paris, alludes to the connection with the theory of equations, and the application which Galois had devised of the theory of groups to the solution of algebraic equations. He suggests that similar considerations in the theory of linear substitutions would, when Riemann's interpretation is employed, open out illuminating geometrical relations on the sphere connected with the regular solids. Indeed, the linear substitutions, when so interpreted, become rotations of the sphere round its different diameters, combined with reflexions at its diametral planes, so that the vertices of each inscribed regular solid, determining as they do rotations which transform them into themselves, give us binary equations in $x + iy$ with appropriate groups of linear transformations leaving them unaltered. In this way, Klein was able to prove that he had obtained all possible finite groups of linear transformations of $x + iy$ by means of the five ancient Platonic figures, together with the flat disc bounded by a regular polygon as degenerate form of a regular solid. And further, applying Galois's ideas, he was able to determine completely the solution of the quintic equation, which was at that time occupying the attention of several of the best mathematicians of the day.

---

* F. Klein, 'Entwickelung,' vol. 2, p. 96.

† The only son, Otto, is an engineer. Two of the three daughters have children. The third, a pupil of her father's, and a war widow, is a teacher in a gymnasium.

‡ 'Ges. Math. Abh.,' vol. 1, p. 412.

9. From the small Bavarian university of Erlangen, Klein passed in 1875 to the Technische Hochschule at the Bavarian capital, Munich. In Germany, where the different levels of society were kept rigidly distinct by what was termed the "Kastengeist," the level of professors at the technical schools was regarded as lower than that of university professors; but Klein looked at things from a broader standpoint. He says himself that there hovered before his mind's eye the vision of a polytechnic school like those of Paris and Zurich, and, in accepting this "call," he felt that he was taking a great step in advance.* The realisation of his plans was rendered easier by the fact that, at the same moment, Brill was called from Darmstadt to Munich. Brill's experience and thorough knowledge of the working and the best traditions of a Technische Hochschule were invaluable to Klein. He enjoyed the advantage, too, of having Dyck as his assistant and Hurwitz as pupil. The latter was also his chief collaborator in the study of the Elliptic Modular Functions, one of the monuments of Klein's fame.†

The years at Erlangen had been only a preparation; it was at Munich that Klein began to feel his feet; here, to use his own words, he "worked himself through to a real mathematical individuality."‡ But, in doing so, he undermined his health, and, at Leipsic, to which university he was called in 1880 as Professor of Geometry,§ he became seriously ill. His breakdown was probably accelerated by the antagonism he experienced at Leipsic. He was much younger than his colleagues, and they resented his innovating tendencies. In particular, there was opposition to his determination to avail himself of the vaunted German "Lehrfreiheit," and to interpret the word "Geometry" in its widest sense, beginning his lectures with a course on the Geometric Theory of Functions.

10. Klein was now at the acme of his mathematical powers. The theorem, which Klein himself prized highest among his mathematical discoveries, known as the "Grenzkreistheorem" in the theory of automorphic functions, flashed upon his mind suddenly at the seaside, in the small hours of the morning, as he sat, propped up on a sofa, because of his asthma.‖ This was at Easter, 1882. The stimulus to this important extension of his earlier work

---

* "Einen grossen Sprung," Autobiography, p. 19.

† Adolf Hurwitz, by W. H. Young, 'Proc. L.M.S.,' Ser. 2, vol. 20, p. xlviii.

‡ Autobiography, p. 20.

§ *Ibid.*

‖ 'Ges. Math. Abh.,' vol. 3, pp. 584, 627.

on the icosahedron and the modular functions had been given by Poincaré's three notes in the 'Comptes Rendus' of February and March, 1881, on Fuchsian Functions, a name given by him, in honour of Fuchs, to a particular type of what Klein afterwards denominated Automorphic Functions.*

Poincaré was only five years younger than Klein,† but had been relatively retarded by the customary somewhat lengthy stages of the mathematical profession in France. In 1880 and 1881 he was *chargé de cours* at Caen. The complete absence of an adequate library there left him quite unaware that, in the grand theory which his three notes foreshadowed, preliminary steps had already been taken. His chivalrous action, when Klein wrote to apprise him of this omission,‡ was to give Klein's name to the remaining Automorphic Functions. The theory of these functions is now marked for all posterity by the intertwined names of Klein and Poincaré. The correspondence that ensued, published in the 'Acta Mathematica,' vol. 39, shows us how the two mathematicians acted on one another as the foundations of the theory were gradually laid.§ After Klein's breakdown, Poincaré continued his work for some time.

11. Obliged by his illness to give up work for a time, and to reduce it materially in the long run, Klein undertook the publication of the one book‖ which he produced single-handed, 'The Icosahedron.' The subject-matter of this little volume, in so far as it originated in Klein's own brain, had been published, in the papers already referred to, in the 'Mathematische Annalen.' We cannot do better than quote Klein's own summing up of the relation in which the book stands to previous work on the subject :—

"In the book many details have been more exactly discussed, and much has been simplified, in which the works of Gordan and Kiepert were of special use to me. On the other hand, much has been suppressed, especially in the theory of invariants. . . . For the rest, the matter is presented, at least as far as regards the equation of the fifth degree, in historical sequence.

---

* See Poincaré's own account, 'Acta Math.,' vol. 38, pp. 44, 45.

† The impression which might be given by Klein and others, *e.g.*, 'Ges. Math. Abh.,' vol. 3, p. 582, that Poincaré was a very much younger man, and very young at the time, is evidently erroneous.

‡ See Poincaré, 'Math. Ann.,' vol. 20, p. 53. "En ce qui concerne les fonctions Kleinéennes, j'aurais dû commettre une injustice si je leur avais donné un autre nom que le vôtre. C'est M. Schottky qui a découvert la figure qui faisait l'objet de votre lettre, mais c'est vous qui avez : 'ihre principielle Wichtigkeit betont,' comme vous dites à la fin de votre savant travail 'Ueber eindeutige Functionen mit linearen Transformationen in sich.'"

§ Reprinted, 'Ges. Math. Abh.,' vol. 3, pp. 587–621. *See* also Klein's historic account, *ibid.*, pp. 577–586.

‖ Klein's classical paper on "Riemann's Theory" ('Ges. Math. Abh.,' vol. 3, pp. 499–574), was also published as a little book.

Hence, apparently, the impression arose in mathematical circles that the object of my method was to render intuitive the accepted theory of the icosahedron. This way of regarding it does not in the least tally with what I still maintain, looking backward. I believe rather that the real foundation of the previous researches of Hermite, of Kronecker, and of Brioschi is only to be found in a preliminary study of the theory of the icosahedron, and in the theory of the elliptic modular functions. The proof of this is that from 1876 to 1880 not only all the points were cleared up which in their theory had remained nebulous; but also, in our rapid advance, questions were attacked which had not even been raised till that moment. I have always looked on those years, in which the momentous steps forward were taken, as the happiest period of my mathematical productivity. Externally, it was characterised by my daily meetings with Gordan. The spot chosen for these meetings was generally Eichstadt, which is half-way between Munich and Erlangen, and we usually spent Sunday there. Gordan used in later years often to speak of the 'Mathesis quercupolitana,' as he called it."*

12. In tracing Klein's career it is striking to contrast it with that of Gordan, the friend just referred to. For eleven years, in spite of his high merit as a mathematician, and in spite of his friendship and collaboration with Clebsch, Gordan had to hold out as Privat-Dozent at the University of Giessen, and only obtained the next step, to an extra-ordinary professorship at Erlangen, a few months before Klein, twelve years his junior, vacated the full chair, which Gordan then occupied till two years before his death in 1912. He seems to have had a rare gift of what we may call mathematical sympathy. In a few words let fall on his 70th birthday, Gordan characterised his ten years' relation to Clebsch as "a happy kind of partnership in which his own task was frequently that of transforming the ideas of that fertile genius."† Noether says‡ of Gordan that no difficulty seemed to him insurmountable and that he threw light on any subject in a truly Socratic manner. This was just the friend for the man of whom Lie said,§ and not without knowledge, "I rank Klein's talents high and shall never forget the ready sympathy with which he always accompanied me in my scientific attempts; but I opine that he does not, for instance, sufficiently distinguish induction from proof, and the introduction of a concept from its exploitation."

13. The moment at which Gordan joined Klein at Erlangen was also opportune. Klein had been more and more progressing from the algebraic and geometric points of view of Plücker and Clebsch to the wider outlook

---

* 'Ges. Math. Abh.,' vol. 2, p. 259.

† Paul Gordan, by Noether, 'Math. Ann.,' vol 75, p. 7.

‡ *Ibid.*

§ Vorrede to the third volume of the 'Transformationsgruppen,' by Lie and Engel (1893).

of Riemann, that of the geometrical Theory of Functions, which in the end dominated his mathematics. Klein had never known Riemann,[*] but he had begun in his Göttingen days to read Riemann's original papers. Gordan, on the other hand, had just met Riemann and sat at his feet for a few weeks in the autumn of 1862, immediately before Riemann's health broke down. The book on Abel's Functions by Clebsch and Gordan had appeared in 1866, three months after Riemann's death, and Klein says of it[†]: "Essentially this important book succeeded in deducing Riemann's results about algebraic curves by the methods of Analytical Geometry. Riemann's own methods were at that time a sort of Arcanum of his immediate pupils, and were regarded by the remaining mathematicians almost with distrust. I can only repeat, on the other hand, what I myself realised at once for curves, that in course of time the progress of mathematics must evidently and inevitably lead to the methods of Riemann himself becoming part of the heritage of every mathematician." If this is the case to-day, and in so far as it is the case, it is chiefly due to the apostleship of Klein. But the chief tenet in Klein's cult of Riemann was only given verbally to his choice disciples, and must remain for ever esoteric: Read Riemann!

It is interesting to note that the turning point in Klein's recovery from his breakdown at Leipsic was marked by his being invited to Baltimore, to take the chair vacated by Silvester in 1884. Indeed, he attributed to this invitation, and to the attractive vistas which it opened to his fancy, the tonification of his mind enabling him to throw off his malady. He eventually declined the invitation, which was not the last he received from an American university. But the days of his great productivity were over, although he had only reached what Dante calls "the middle of our walk in life"; from this time, for many years, he devoted himself to the development of his mathematical school.[‡]

In 1886 he was called to Gottingen. The great city was indeed no home for the true German, and it was with enthusiasm that the Kleins migrated to "the little garden-town."[§] Time has so touched with the pencil of prosperity the features of Göttingen, that to-day it is difficult for the visitor to realise what it was in the primitive simplicity of those days. Göttingen was then still a medieval poem. It was here that the happiest home-days of the family were passed, in the comfortable villa, all their own, built by themselves, in its cool shady garden. On one side it was close to the Auditorium, on the other, near by, was the magnificent forest, the Göttingerwald, stretching to the Harz Mountains. Every day, after midday dinner and a short siesta, Klein used to walk up into the forest in mathematical converse with his colleagues and friends, after which he received his special students upstairs in his book-lined* study, and discussed their work, or he prepared his own lectures.

---

[*] 'Ges. Math. Abh.,' vol. 3, p. 477.

[†] *Ibid.*, p. 490.

[‡] Autobiography, p. 22. It was in the excellent historical seminar of Waitz at Göttingen in 1868–9 that Klein first got the idea of the importance of a well-organised department. Autobiography, p. 15.

[§] *Ibid.*

16. Göttingen did not disappoint Klein and he refused every "call" elsewhere. He saw the university raised once more to a world-celebrity which even in the days of Gauss had never been surpassed, and to a material magnificence which had never been even imagined. The turning-point came when, as he himself humorously pointed out at a Christmas gathering of his students in 1893, the centre of the universe veered round to Gottingen, following as it always must, the advent of Eve's daughters. In fact, Prussian university life had been gently invaded. Two American women, inspired by Klein's visit to Chicago that year, had followed him to Europe, and a Girton girl had appeared at the same time independently†; it was she who became eighteen months later the first female Doctor of Philosophy at a Prussian University.‡

17. At the Auditorium, on the top floor, was Klein's lecture room; in the days of his best lectures it rarely contained more than a dozen hearers, and these nearly all foreigners,§ but at the beginning of the present century there were near on a hundred students, mostly Germans, in his audience. Hard by was the wonderful mathematical reading-room, with the adjacent room lined with glass-cases full of models; here the professor held, after his lecture, his "Sprechstunde," when any of his pupils could come to him with requests or questions. For a very small fee the student received a pass-key enabling him to enter the reading-room at any hour and avail himself of the books, which, unlike those at the University Library, might not be taken out, and of the manuscript and lithographed notes of lectures; in particular he could consult the written "Ausarbeitung" of Klein's lecture of the day before, made by Klein's assistant and previously submitted to himself, it happened, not infrequently, that rash statements of the professor were in this way controlled and corrected, these assistants, particularly the too early lost Ernst Ritter, being for the most part mathematical collaborators of a high order.‖ All this, which had been organised by Klein, has now been developed into a Mathematical Institute in a separate building.

18. We have here touched only on a very small portion of Klein's activity in the improvement of the university institutions and arrangements. Incidentally

---

* Klein's magnificent collection of books and papers is now at the University of Jerusalem.

† Autobiography, p. 28, is here misleading.

‡ 'Ges. Math. Abh.,' vol. 3, Anhang, Dissertations No. 38.

§ Here Professor Courant's 'Gedächtnisrede' might give a false impression.

‖ 'Ges. Math. Abh.,' vol. 3, Anhang, p. 14.

this was facilitated by the part he played in bridging over the cleft which existed between scientific and industrial circles. The brilliant idea of the "Göttinger Vereinigung zur Forderung der Angewandten Mathematik" was Klein's own, and was suggested to his mind by his observations in America. This society, founded in 1898, included some of the foremost members from among the manufacturers. The funds which these public-spirited individuals placed at the disposal of the university served to eke out the generous, but necessarily inadequate, contributions of the Prussian Government. In this way new Institutes for Geodesy, Astronomy and Assurance, technical Mechanics and technical Electricity were created, and corresponding chairs founded and filled.*

19. What the university and the town of Göttingen owe to Klein can never be forgotten; the monument is the Göttingen of to-day.† Perhaps the most important influence which he directly exerted—putting aside his constant advocacy of the rights of Applied Mathematics on an equality with Pure—was in the election of his colleagues. A single appointment to a professorial chair based on reasons, openly or tacitly accepted, other than that of *efficiency*, is not only sufficient to undermine the prestige of a university but also to corrupt its ideals for a considerable period of time. Klein's two principles in recommending the call of a new professor were, first, *individual pre-eminence*, and second, *collective representativeness*. In particular, it was Klein who brought Hilbert, and indirectly also Minkowski, to Göttingen, for Hilbert, on the occasion of his refusing a call to Berlin, made it a condition of his remaining at Göttingen that a place should be found there for Minkowski: on similar occasions when Klein refused such calls, he always used the opportunity to obtain concessions in the interest of the university.‡ This point deserves our attention; it is habitual in German university life, and is one of its noblest traits, without anything corresponding to it in our own universities.

20. It would be difficult to gauge how far Klein's influence in his own country was enhanced by his entry into the Prussian Herrenhaus in 1907, chosen by the University of Göttingen as its representative. This election had been partly prompted by the desire to see him expend some of his unsleeping energy outside the university. However, though he sat on various committees connected with education, he hardly ever spoke.

21. To mathematicians for generations to come, Klein is likely to be thought of as the principal promoter of the Enzyklopädie der Mathematischen Wissenschaften. He had a remarkable memory, and never forgot what he had once imbibed nor where he had come across it; but he had no unusual facility at any

* 'Ges. Math. Abh.,' vol. 2, pp. 507 et seq.

† See Courant's 'Gedächtnisrede.'

‡ Autobiography, p. 32.

time of his life for mastering the contents of printed books,* and, as time went by, he found himself mathematically stranded, and quite incapable of following, even from afar, the actual course mathematics had taken, except where, as in the Theory of Relativity, there was a distinct connexion with the realms in which his own early successes had been achieved; and, even then, it was only to the extent of this connexion that he was able to follow. The Enzyklopädie was, from Klein's point of view, an effort to render accessible to his pupils, to himself, and to the mathematical public at large, the bulk of existing mathematics. One day in the '90's the concept of the Enzyklopädie was formulated by Klein in the presence of the writer: the progress of mathematics, he said, using a favourite metaphor,† was like the erection of a great tower; sometimes the growth in height is evident, sometimes it remains apparently stationary; those are the periods of general revision, when the advance, though invisible from the outside, is still real, consisting in underpinning and strengthening. And he suggested that such was the then period. What we want, he concluded, is a general view of the state of the edifice as it exists at present.

Klein undertook, for his own part, in accordance with his pronounced taste for Applied mathematics, the editorship of the volume on Mechanics; but over and above this, he was indeed the moving spirit of the whole undertaking. He says in his Autobiography (p. 31):—" I always emphasised above all things the view that, in the Enzyklopädie, Applied Mathematics must be treated on an equal footing with Pure. But in this connection there were very considerable difficulties to be overcome. Our object was not so much to publish articles illustrating the ordinary routine in the application of mathematics to each different domain, but, rather, by searching discussion with the experts, to get at the mathematical kernel of the subject. With this in view, I was entrusted with the task of obtaining the necessary groundwork for the production of the applied volumes (Mechanics, Physics, Geophysics and Astronomy) by personal contact of a very varied character. In particular this demanded on my part journeys to Austria, Italy, France, England, Holland and Denmark"

It was on one of these journeys that he arrived, on his 50th birthday, at Turin, and was fêted by the mathematicians of that city, where the present writer and his wife—Klein's favourite pupil, as the saying went—were then studying. Placed at dinner next to the lady, he whispered, in answer to her felicitations:—" Ah! I envy you. You are in the happy age of productivity. When everyone begins to speak well of you, you are on the downward road."

The conception of the Enzyklopädie was grandiose, and its mere existence had amazing effects. We can understand this when we reflect that, in England, in 1892, the Theory of Functions of a Complex Variable, due to Cauchy—not to speak of Riemann—had barely been brought within the borderland of knowledge. Yet the Enzyklopädic as a whole cannot be regarded as a success. Collaboration between writers of different articles, living in different lands and writing on subjects apparently widely apart, was impracticable, even when the underlying mathematical scheme was similar. The want of unity of design had the effect of making the construction expand to an inordinate degree, and even so, the articles themselves were sometimes felt to be insufficiently detailed, and had to be supplemented by reports in the 'Jahresbericht der Deutschen Mathematikervereinigung.' We cannot help recalling the claim of Roger Bacon, that he would undertake to explain to a person of ordinary intelligence in three or six months, all that he himself knew, and this, as is almost certain exceeded the accumulated learning of his time. That is the feeling which every competent mathematician must *mutatis mutandis* have with respect to our existing mathematical science. All that is essential should be explainable in a small number of volumes of moderate size. But we shall attain this ideal only by a series of approximations, and the Enzyklopädie, if the first, seems a very rough one indeed. It would appear to be characteristic of the Germanic mentality to generate wide-embracing conceptions and, with admirable perseverance, to carry them through to some sort of realisation, while the Romance mind would hesitate to undertake anything unless it was sure of attaining excellence in form.

---

* In particular, see 'Ges. Math. Abh.,' vol. 1, p. 241, where he explains how he gathered his knowledge of v. Staudt's work.

† Op., *ibid.*, vol. 2, p. 510.

22. Klein's appetite for work was such that, long before there was any prospect of the Enzyklopädie being completed, finding he was obliged to leave much of the organisation in other hands, he had turned to the international question of mathematical teaching in the schools. He turned perhaps to something more concrete, which brought him into contact with minds other than those of the mathematical expert, something more comprehensible in fact than the huge colossus which was threatening to remain nothing but a torso. In the matter of the International Mathematical Teaching Commission Klein, as President, formed one of a triumvirate, of which another, the pioneer of the movement, was a world-travelled American, David Eugene Smith, and the third, as General Secretary, a native Swiss professor, Henri Fehr. This international work was checked and rendered abortive by the Great War.

23. It is impossible even to touch on all Klein's schemes, yet we cannot but allude to a vast undertaking, still uncompleted, in which he was the prime mover, the editing and issuing of Gauss's Works. For Klein, Gauss represented the universal in Mathematics, and in this connexion he held Gauss before himself as the master to be copied.* But, if Gauss was here the prototype, it was buried in his individual person; the apostle of universality in Mathematics

---

* There was hardly a new mathematical idea formulated, but Klein would cite a passage from Gauss in which it had been foreshadowed.

was Klein, and, without hyperbole, he may be termed the founder of what the 20th Century means by "Mathematical Science," raised equally above departmental specialism and narrow nationalism. Indeed, during a period when the forces of material progress were, at first secretly, but afterwards openly, undermining the basis of the unity of civilised nations, generating jealousy, and inciting to war, Klein embodied the contrary principle, and, by his great power in his own country, and his wide-spread influence in the world of intellect, constituted one of the greatest forces tending to international amity and peace. In particular, it was he who revived the practice—prevalent in the Middle Ages, when Latin was the universal tongue of Learning—of inviting distinguished men from other countries to come and give lectures.*

24. In looking through Klein's Autobiography, one is struck by his constant references to his friendships. It was characteristic of him that they all bore directly, or indirectly, on his mathematical development and on his power of organisation. He was indeed a man without a hobby; in particular, and this is curious and interesting from a psychological point of view, although a German, and although endowed with an excessive acuteness of hearing,† he could not distinguish one tune from another.

Social intercourse for Klein meant the interchange of ideas, and for that he was as eager as an ancient Athenian. It was in such give-and-take that his own conceptions took form. There is, perhaps, no contradiction in saying that Klein was never the originator of his own ideas. He had not the generating force of a Cauchy or a Georg Cantor, but he had a phenomenal power of grasping the import of a suggestion, and working it out on a grander scale than any before him had imagined.

We have already mentioned incidentally several of those close friendships that directly influenced and even inspired Klein as a mathematician. The most striking of all, his friendship with Lie, merits a place by itself. It takes us back to his stay at Berlin in 1869. Klein writes‡ :—" The most important event of my time in Berlin was certainly that, towards the end of October, at a meeting of the Berlin Mathematical Society, I made the acquaintance of the Norwegian, Sophus Lie. We had, in our work, been led from different points of view finally to the same questions, or, at least, to kindred ones. Thus it came about that we met every day and kept up an animated exchange of ideas. Our intimacy was all the closer, because at first we found very little interest for our geometrical interests in our immediate neighbourhood." At this time it was chiefly with Line Geometry that the two were occupied.

---

* In particular Poincaré came to Göttingen in 1894 or 5, and gave a lecture in French to the mathematical professors and students.

† This enabled him to catch the faintest whisper during his lectures. Also, commonly, when away from home, he could not go to sleep unless he stopped up his ears with cotton-wool; this was the case, for instance, during his short stay at Trinity College in 1897, on the occasion of his receiving his honorary degree at Cambridge.

‡ 'Ges. Math. Abh.,' vol. 1, p. 50.

In the summer of 1870 Klein and Lie went together to Paris, where they had rooms side by side, and lived in the closest bonds of friendship. Their intention was to go on to England in the winter, but the outbreak of the Franco-German War prevented it. In this connection Klein writes:—"This striving after the greatest possible breadth of scientific outlook, for which an acquaintance with what was being done in foreign countries seemed important, was very little understood at that time in Germany. Thus, for instance, when, urged by my father, I tried to obtain letters of recommendation from the Minister of Education in Berlin, I received the official answer:—We have no use for French or English mathematics."*

In Paris the two friends did not attend lectures, but worked together. It was, however, undoubtedly by the personal contact with the younger French mathematicians, particularly Camille Jordan, that the attention of the pair was directed to Galois's theory of groups. Jordan's treatise on the 'Theory of Substitutions' had just appeared, and was to the two young men "a book with seven seals."† We continue in Klein's own words:—"It was, however, with Darboux that we were most intimate. At that time the French had been occupied with new investigations in metrical geometry (inversion, constant use of the circle at infinity), and this was quite unknown in Germany. Now it appeared that these were most closely related to our own work in line geometry. . . . For Lie, what was perhaps most interesting was the researches of the French into the geometrical theory of differential equations. It was by the coalescence of the two circles of ideas that Lie came to discover the connection between lines of curvature and minimal lines on a surface."†

And here Klein gives a vivid sketch from this period of his life:—"I had risen early one morning at the beginning of July, 1870, and I was just going out, when Lie, who was still in bed, called me into his room, and began explaining to me what he had found out during the night, the connexion between lines of curvature and minimal lines. I did not understand one word. Anyway, he assured me that it followed that the minimal lines on Kummer's surface must be algebraic curves of the 16th order. In the course of the morning it flashed upon me, while I was seeing over the Conservatoire des Arts et Métiers, that these must be the same curves of the 16th order which had already appeared in my theory of the Line-complexes of the first and second degree, and I succeeded at once in carrying out the geometrical proof, independently of Lie's transformation. When I got back in the afternoon, at 4 o'clock, Lie had gone out, and I left a letter for him containing my proof."‡

---

* Autobiography, p. 16.

† 'Ges. Math. Abh.,' vol. 1, p. 51.

‡ 'Ges. Math. Abh.,' vol. 1, p. 97.

25. The outbreak of the war, Klein's consequent hasty departure, leaving his Norwegian friend behind, then the continuation of their mathematical conversations by letter, have an interest beyond that of the immediate consequences to Lie. While Lie remained at Paris, all went well; but when, in the heat of August, he started for Italy on foot, he was arrested at Fontainebleau as a spy, and kept in durance for four weeks, till, in fact, their friend Darboux had succeeded in explaining the inoffensive mathematical nature of the mysterious letters, in German, and in Klein's notoriously illegible handwriting, found in Lie's possession.

26. This episode was one of the manifestations of the unbroken continuance during the Franco-German war of friendly scientific relations between himself and his newly-won French friends, which made such an indelible impression on Klein's mind. We may say that those were still the days of scientific chivalry. When the Great War broke out in 1914, Klein, like most of us, was incapable of conceiving any other state of affairs, but he was, as he had in fact been in 1870, carried away by the wave of patriotic fervour of the moment. When he therefore began to perceive that there was serious danger of what he regarded as a purely political cataclysm affecting international scientific relations, he lost, for one unfortunate moment, that dignified poise of judgment which was one of his most striking characteristics. Invited by telephone to sign a document, which, according to the impression left on his mind* was to make an appeal to the scientists of Europe to maintain an objective attitude during the struggle, he impetuously acceded to the request. He did not see the text of the document until it appeared in the daily papers with his signature, as one of the too well known 93, printed below. It has been very damaging to Klein's international position that he permitted his name to appear without protest at the foot of a document whose tenour was so entirely antagonistic to the whole tone of his life. And he was too proud, even after the Armistice, to withdraw publicly his signature. He desired his friends to know, and to tell, how it had been obtained, but he felt that, for himself, he stood upon his whole career, and not upon an isolated and misinterpreted action.†

---

* All this is founded on an autograph letter to my wife dated December 7, 1918. See her Obituary notice in the *Times* of July 9, 1925. Ultimately, Klein allowed an extract from this letter to be printed in a Berlin daily paper, together with explanations from most of the 93 of more or less the same tenour.

† In a letter of July 15, 1919, to the same, the conclusion is as follows:—"Allgemein verbietet mir den Sinn für Objektivität mich über Dinge zu aeussern, die ich, wenn überhaupt, nur aus den subjektiven und einander wiedersprechenden Aeusserungen irgend welcher Zeitungen konne. Also bleibt für mich, was ich die ganze Kriegszeit über getan habe: Schweigen und Arbeiten. Für die wenigen Jahre, die ich noch vor mir habe, werde ich damit auskommen.

"Die Welt aber wird ihren Lauf nehmen und die Völker werden sich eines Tages wieder zusammen finden. Vorläufig ist eingetreten, was einst beim Turmbau von Babel der Fall war, sie verstehen einander nicht mehr."

27. Lie was seven years older than Klein, but their standing in mathematical productivity was about the same. Lie had only begun mathematics seriously at 26,[*] whereas Klein, at the age of 20, when they first met, had been for three years recognised as a mathematician and a writer. Lie's first publication in which he represents the ∞^4 imaginary lines of the plane by the ∞^4 real lines of space, had been published in 1869, and Klein's Doctor's Dissertation "on the Transformation into a Canonical Form of the General Equation of the Second Degree between Line-co-ordinates" was presented to the University of Bonn in December, 1868. Both works were stimulated by the ideas of Plücker; indeed Engel, who was intimate with Lie, attributes to him a phrase in which Lie calls himself a pupil of Plücker, which he was not. But the printed phrase to which Engel gives the reference,[†] bears the stamp of being written carelessly, and is in a language other than Lie's native tongue; it hardly bears the interpretation Engel puts on it. When Lie writes "my extensions are quite obvious for a pupil of Plücker's," one cannot help thinking that the pupil of Plücker's referred to in general terms, if specialised at all, was not Lie, but Klein. Indeed it is difficult to believe that Lie, the Norwegian, would have gone as far as Engel, typically German, in attributing even the inspiration of his own ideas to a German. Lie was excessively tenacious of his own rights to his own ideas, and resented, as is well-known, any suggestion that they were in any sense second-hand. He quarrelled with Klein, his early friend, on this very point, and in the introduction to the third volume of the 'Transformationsgruppen'[‡] he repudiates the notion that he was a pupil of Klein's; "rather," he says, "the contrary was the case." This irritated Klein, and justly. His generous nature was incapable of making the former of these suggestions, and his frank and healthy self-confidence prevented his accepting the latter of them. It is so rare in this world of ours that misunderstandings are ultimately cleared up, that it is a real pleasure to know that this was the case with Klein and Lie. In the vivid picture which Klein gives in his Autobiography and in the Collected Mathematical Papers of the interweaving of his own work with that of Lie, there is no bitterness. I am so fortunate as to be able to add to what may be there read a little sketch by the hand of Mrs. Klein,[§] sent specially for the purpose of the present writing.

"The relation to Lie was an intimate friendship in mathematical and in personal respects. That remained the same also later, when Lie visited us in Leipzig. And I, too, was fond of him, the powerful Northman with the frank open glance and the merry laugh, a hero in whose presence the common and the mean could not venture to show themselves.

---

[*] "Sophus Lie," by F. Engel, 'Jahresber. d. deutschen Math. Vereinigung,' vol. 8, p. 30 (1900).

[†] 'Math. Ann.,' vol. 9, p. 246; 'Engel,' p. 34.

[‡] *Loc. cit. supra.*

[§] Dated August 15, 1925. She died last year.

"Then he became my husband's successor at Leipzig, and there he was seized with home-sickness. How well I understand that! He, the free, accustomed to his rough but beautiful northern home, how could he stand the great smoky city, the high houses, the narrow streets, the puny Saxon people. He suffered from melancholia, he was taken to a sanatorium, but there things only became worse, for they robbed him of his work and of his freedom. There, in his embittered state of mind, it must have been that he wrote the malicious things about my husband which were both painful and incomprehensible to him.

"But soon he understood that his best friend was ill and that he could not be held responsible for his actions. It was not only magnanimous and good, but also wise of my husband, not to enter into any polemic, but to let the thing simply rest. And he was not mistaken in his friend.

"One summer evening, as we came home from an excursion, there, in front of our door, sat a pale sick man. 'Lie!' we cried, in joyful surprise. The two friends shook hands, looked into one another's eyes, all that had passed since their last meeting was forgotten. Lie stayed with us one day, the dear old friend, and yet changed. I cannot think of him and of his tragic fate without emotion. Soon after, he died, but not before the great mathematician had been received in Norway like a king."

We will only add that, physically, Klein was very tall, erect and slim, with rich brown wavy hair and characteristically sparkling light blue eyes, with a genial glance which has not been completely caught in the photograph here produced.

W. H. Y.

HENDRIK ANTOON LORENTZ—1853-1928.

Hendrik Antoon Lorentz, one of the greatest scientific figures of our time, was born at the small town of Arnhem in Holland on July 18, 1853, and died at Haarlem on Saturday, February 4, 1928.

He attended Mr. Timmer's Primary School in Arnhem until he was thirteen, when he entered the newly-established High School. At the age of seventeen he passed the examination for admission to the University of Leyden. Two years later he returned to Arnhem to be a teacher in a public evening school and studied alone for his doctorate. He took this degree at the University of Leyden at the age of twenty-two, the subject of his thesis being 'The Theory of the Reflection and Refraction of Light.'

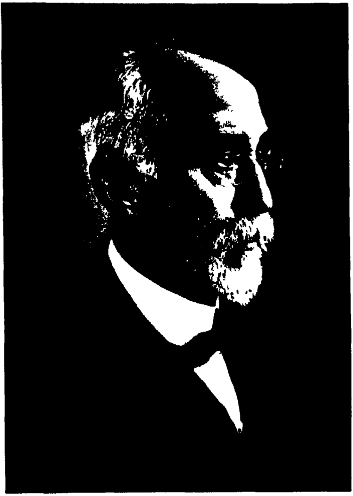
This was the most prominent outstanding problem of the electromagnetic theory of light which Maxwell had left unsolved. The natural way of attacking it was to follow the acoustical analogy and satisfy Green's conditions of the continuity of the displacements and tensions at the interface. There was, however, some latitude as to the transcription of the electromagnetic quantities into their mechanical equivalents. On Fresnel's formulation, at least, it appears that there are more conditions at the interface than can be satisfied without the inclusion of longitudinal waves which are precluded in optics.

Maxwell never made up his mind on this question. Lorentz resolved it by showing that the correct conditions to be satisfied at the boundary were the electromagnetic conditions of the continuities of the tangential components of the electric and magnetic intensities and the normal components of the electric and magnetic inductions respectively. Apart from Maxwell, he was largely influenced in this work by the writings of Helmholtz.

The power and importance of this dissertation was immediately recognised, and two years later Lorentz was called back to Leyden as Professor of Theoretical Physics in the University at the early age of twenty-four.

In 1892 Lorentz put forward and vindicated the theory of electrons in a paper ("La théorie électromagnétique de Maxwell et son application aux corps mouvants," 'Arch. Neérl.,' vol. 25, pp. 363-551), which both immediately and subsequently exerted a most profound influence on the development of theoretical physics.

Up to the time of Faraday and Maxwell all physical theories having any pretensions to fundamentality had been built on the application of Newtonian mechanics to the interaction of material particles. It was implicitly, if not explicitly, assumed that all physical phenomena would ultimately be found explicable in such terms provided the initial conditions could be adequately specified. This supposition, originally restricted to mechanics, where its



H. A. Lorentz

(*By permission, from ‘Nature,’ vol. cxi*)

consequences had been abundantly verified, gradually extended throughout the whole of physics with the firm establishment and generalisation of the principles of the conservation of energy and momentum. Faraday had, however, found that the phenomena of electrostatic and magnetic induction were not to be accounted for in terms of the more distant elements of electric charge or of magnetism, but that the properties of the intervening media had a profound result on the phenomena observed experimentally. For Faraday and Maxwell the field was just as important as the charge elements; Maxwell's 'Electromagnetic Theory of Light' was built fundamentally on it. In spite of its success in the equality of the ratio of the units with the velocity of light, the electromagnetic theory of light as left by Maxwell had some serious imperfections which his successors were not immediately able to remove. One of these had been effectively dealt with by Lorentz in his dissertation. Another lay in the fact that the coefficients which expressed the electric and magnetic quality of the medium were not constants, as the theory assumed, but depended on such variables as the frequency of the light. Thus the theory had not developed to the extent of accounting for dispersion, although Maxwell himself must have seen how to overcome this difficulty in a general way, since he published the essential features of the solution in the form of an examination question. A still more serious difficulty was met with in attempting to apply the theory to moving media. Both Maxwell and also Hertz, who subsequently attacked this problem, had come to the conclusion that the light in a moving medium should be convected with the full velocity of the medium, whereas it was known that this velocity was only shared to a limited extent; the well-known formula of Fresnel had been well established experimentally.

It is the supreme merit of this work of Lorentz that he saw clearly the minimum assumptions which accounted for the essential facts and that he carried those assumptions through to their inevitable logical conclusions. His theory was not of the supple kind which by slight modifications here and there can be made to accommodate inconvenient facts overlooked or undiscovered at the time of its development. It was of the kind which must stand or fall by the truth or falsity of the fundamental assumptions on which it is built. Every subsequent investigation has only served to confirm its entire validity within the field of the classical physics. Within this range its applicability is complete, exact and unambiguous. It must thus be regarded as the culmination of the classical philosophy, and, so far as can reasonably be foreseen, it will always maintain this position.

The essential elements of Lorentz's theory which distinguish it from those of Maxwell and his other predecessors are the attribution to electricity of the properties of atomicity and universality in the composition of matter and the reduction of the various media to a single one, the aether. This aether, however, was quite different from the aethers of previous optical theories. It was always at rest and unaffected by the motion of material bodies through it. It is in fact hardly unfair to regard it merely as space, in which the electric and magnetic vectors could interact in accordance with Maxwell's specification.

Probably none of these conceptions were exactly new in 1892, but the synthesis at any rate was novel. From the beginning of his career Lorentz had been attracted by the hypothesis of small charged particles embedded in matter as the real seat of the reaction of the field intensities (see for example, 'Verh. d. Akad. van Wet.,' vol. 18 (1879)). This was, of course, not the first electron theory, the older electrodynamical theories of Weber, Riemann and Clausius implied the electronic feature in some form or another, but in these older theories the electrons acted directly on each other by forces at a distance In Lorentz's theory the electrons only acted on each other through the aether in which they were embedded. This enabled Lorentz to adopt Maxwell's analysis for the cases in which it was known to be valid, that is to say, cases which did not include motional effects, refraction or dispersion, and he was thus led to the well-known field equations

$$\mathrm{div}\,\mathrm{E} = \rho,\ \ \mathrm{div}\,\mathrm{H} = 0,\ \ \mathrm{rot}\,\mathrm{E} = -\frac{1}{c}\dot{\mathrm{H}},\ \ \mathrm{rot}\,\mathrm{H} = \frac{1}{c}(\dot{\mathrm{E}} + \rho\mathrm{V})$$

To find the mechanical force on a current element in an electromagnetic field, he considered the case of a single charged particle moving in an atmosphere of similar particles; this led to the additional equation

$$\mathrm{F} = e\mathrm{E} + \frac{e}{c}(\mathrm{VH}).$$

These equations are satisfied by the field components or by a single electron in the free aether. To apply them to material media they must be averaged over the electrons present in the media in a manner appropriate to the conditions of the problem in hand. In considering the force on an electron in the interior of a dielectric placed in an electric field it is necessary to consider not only the force due to the actual charges on bodies at some distance, but also the force due to the doublets induced by the force acting on the electrons which help to constitute the medium. This leads directly to the idea of the polarisation P and it is found that the so-called fictitious charges of the old polarisation theory of Poisson represent real charges due to an accumulation of displaced electrons in this theory. After this process of averaging, the equations become

$$\mathrm{div}\,\mathrm{D} = \rho,\ \ \mathrm{div}\,\mathrm{B} = 0,\ \ \mathrm{rot}\,\mathrm{E} = -\frac{1}{c}\mathrm{B},\ \ \mathrm{rot}\,\mathrm{H} = \frac{1}{c}i,$$

where D and B are the inductions and i represents the sum of the conduction, convection and displacement currents + rot ($\mathrm{P}w$). This last term was substituted by Lorentz for a term rot ($\mathrm{D}w$) which Hertz had added to Maxwell's equations; its correctness was established by the various experiments of Roentgen, Rowland and Eichenwald.

In Lorentz's hands this theory led directly and inevitably to the formula of Fresnel, to the accuracy of the first power of the aberrational constant, for the effect of the velocity of a moving dielectric on the speed of light travelling through it, as well as to the formulæ of the classical theory of dispersion. The former result is almost obvious from the fact that the motional effects are due to the induced fictitious polarisation charges which are proportional to the electric intensity multiplied by the difference between the square of the refractive index and unity. The displacement of the polarisation electrons in a periodic field of force must depend, in a manner well understood from the theory of forced vibrations, on the relation between the frequency of the field and the natural frequency of the electrons. These considerations lead at once to the required connection between the frequency of the field and the magnitude of the refractive index or dielectric constant. Lorentz's result that the motional effects are proportional to $\mu^2 - 1$ (or $K - 1$), as opposed to a previous result of Hertz which made them proportional to the dielectric constant itself, was confirmed by specific tests made on charged condensers moving in a magnetic field by Blondlot and H. A. Wilson respectively.

There still remained one outstanding difficulty in the way of a stagnant aether, the negative result of the Michelson and Morley experiment. In an earlier paper Lorentz had proposed to overcome this by a compromise between the older theories of Fresnel and Stokes, an explanation rather forced and somewhat complicated. In 1892 Fitzgerald put forward the hypothesis that all moving bodies contracted their dimensions in the direction of motion in the ratio $(1 - v^2/c^2)^{\frac{1}{2}}$ to unity. This hypothesis which accounted for the negative result of the Michelson and Morley experiment was immediately adopted by Lorentz, who had no difficulty in making it harmonise with his general point of view.

The position at that time, as it is to-day and probably always will be, was that every experiment designed to demonstrate the effect of moving matter on electrical and optical phenomena led to a negative result unless relative motion of different material parts of the system was involved. The motion of the system relative to the aether or anything else made no difference provided the velocity was uniform for all its parts. The significance of this did not escape Lorentz and he proceeded to investigate what changes were necessary in the fundamental variables of the electromagnetic field, both dependent and independent, for the validity of this result to be general. In 1895 ('Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern') he succeeded in showing that the fundamental equations retained their form unaltered in a transformation to axes moving with uniform velocity v parallel, let us say, to the x axis provided the time t' in the moving system was changed from t to a local time $t' = t - vx'/c^2$. The values to be taken for the dependent variables, the field intensities in the moving system, were those naturally required as a consequence of the views he had previously developed.

Although this result was only established to the accuracy of the first power of v/c, this was sufficient to cover all well-established cases except the Michelson and Morley experiment. By 1900 Larmor ('Aether and Matter') had succeeded in proving the theorem to be accurate to the order of v^2/c^2, and this accounted for the Michelson and Morley result and the Fitzgerald contraction. In 1903 ('Proc. Amst. Acad.,' vol. 6, p. 809) Lorentz showed further that for all values of v less than the velocity of light, electromagnetic and optical events in a moving system were independent of the (uniform) velocity of that system. This result is known as Lorentz's Principle of Correlation. It originated as a consequence of the conflict between Lorentz's fundamental theory of 1892 and the results of Michelson and Morley and of other types of experiment devised to test the same or similar questions. Its content is the same as that of the Restricted Principle of Relativity which it preceded and of which it must be regarded as an alternative and equivalent statement.

On Lorentz's theory the emission of light and heat by bodies was an immediate consequence of the vibrations and motions of the charged particles of which they were composed. The effects to be expected from any type of motion required in general elaborate analysis. Inasmuch, however, as the charged particles must normally be in some configuration of equilibrium in order to preserve the permanence of the material substance it was to be anticipated that a very common occurrence would be the execution of small oscillations about an equilibrium configuration. The effect of such oscillations would be the emission of light of the same frequency as the natural frequency of the oscillators. This was merely an application on the subatomic scale of Maxwell's principles which had already been carried out experimentally on the macroscopic scale by Hertz.

In 1896 Zeeman discovered that the frequencies of spectral lines were altered when the emission occurred in a powerful external magnetic field. The existence of a connection of precisely this kind had been a conviction of Faraday's, and both he and Tait had looked for it in the laboratory, but the fields at their disposal were inadequate. The theoretical explanation of the simplest type of this effect was immediately produced by Lorentz ('Phil. Mag.,' vol. 43, p. 232 (1897)). An examination of the data showed that the sources of emission of the light were negatively charged particles whose mass per unit charge had the same value as that for the carriers of the cathode rays whose true nature was being elucidated about that time by the work of J. J. Thomson. The Zeeman effect thus at once furnished the proof of the general existence in matter of a subatomic electrical atom, the electron. It had, of course, many other important consequences and has furnished what is probably the most powerful tool in the hands of the spectroscopic investigator. The importance of this discovery and its implications was immediately recognised and was signalised by the joint award to Lorentz and Zeeman of the Nobel prize for Physics in 1902.

The foregoing represents a mere outline of some of Lorentz's more important achievements. As Larmor has truly said: a survey of his whole life's work is a liberal education in the history of physical science during the past half century. There is hardly any fundamental question which his writings have not materially illuminated. One of the matters which engaged much of his attention was the emission of black body radiation as a function of temperature. His first method of attacking this problem was to analyse the radiation emitted during the motion of electrons in metals or any similar structure. He supplemented this later by using other methods. But in agreement with other investigators he always came to the same conclusion, namely, that provided the motions were to be fundamentally governed by the Newtonian dynamical principles there was no escape from the Rayleigh-Jeans formula which was in conflict with experiment. Another notable achievement was that of the calculation of the thermal and electrical properties of metals on the classical theory of metallic conductors ('Theory of Electrons,' p. 266). Among the other subjects which also claimed his attention may be specially mentioned the width of spectral lines, the theory of gravitation and the principle of relativity.

An early investigation of Lorentz dealt with the refractive index and specific inductive capacities of bodies as a function of the density. In this he arrived at a well-known formula sometimes called the Lorentz-Lorenz formula because it was also put forward independently and on different grounds by the Danish physicist, L. V. Lorenz. Whilst this formula is derived from Lorentz's electron theory it is not an inevitable consequence of that theory but involves the use of assumptions which are not essential to it. At this point the detailed structure of the individual molecules comes into play and the assumed conditions are probably too much simplified for it to be exactly applicable to real substances. Nevertheless it is correct in principle and is satisfactory as a first approximation. It is very curious that L. V. Lorenz should have anticipated very much earlier the one considerable omission from Lorentz's earlier formulations of his electron theory, namely, the introduction of the retarded potentials, which are vital in any case of accelerated motion. In point of fact Larmor had the earlier success in dealing with accelerated electrons ('Phil. Mag.,' vol. 14, p. 503 (1897)).

After holding the Leyden Professorship for 35 years, Lorentz in 1912 accepted an invitation to go to Haarlem as curator of the laboratory of the Teyler Institution. This relieved him of the more arduous part of his duties as

Professor at the University of Leyden, but right to the end of his long and active career he continued to go to Leyden once a week to deliver a lecture and to discuss outstanding problems with the students.

Apart from his papers in the scientific journals, Lorentz has enriched us with a large number of publications of more extended range. In addition to the 'Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Korpern' and the 'Theory of Electrons,' of which the essential content as well as the treatment is highly original, these include a number of surveys of various extended branches of physical knowledge. As a rule these were the fruit of some course of lectures delivered under some special conditions. They are of permanent value not only on account of the independence of the author's treatment but also on account of the accuracy and clearness of the exposition. He also found time to publish two textbooks which have been much appreciated on the Continent.

He was much concerned with the important national undertaking of draining the Zuider Zee, and after the war he devoted a great deal of attention to fostering international scientific relations, especially between the late belligerents.

One of the great events of Lorentz's life was the celebration on the 11th of December, 1900, at Leyden of the twenty-fifth anniversary of his doctorate. This was attended by many leading scientists from various countries who contributed a volume of memoirs in his honour. In 1926 the Senate of the University of Leyden made him an honorary Doctor of Medicine as a mark of appreciation of the trouble he had given himself in regard to the instruction of medical students. In the same year physicists all over the world commemorated the fiftieth anniversary of his doctorate by subscribing a fund for the creation of the Lorentz Foundation for the promotion of theoretical physics and of international intercourse among young physicists The object of this was to send young men from the Netherlands to other countries and to invite young foreigners to the Dutch Universities. There were thousands of subscribers, large numbers of whom were not scientists.

That Lorentz was a man of remarkable intellectual powers is obvious from his writings, but he was pre-eminently one of those beings of whom a full appreciation is only to be obtained by personal contact. Although steeped in his own investigation of the moment, he always seemed to have in his immediate grasp its ramifications into every other corner of the universe. To those who knew him he gave the impression of having the power of bringing more of nature into focus in his mental vision at one time than any of his contemporaries. The singular clearness of his writings provides a striking reflection of his wonderful powers in this respect.

He was gifted with a manner of indescribable charm and a surprising modesty which was most attractive. In spite of concentration on the matters which were uppermost in his own mind he always seemed able to take an interest in the affairs, whatever their nature might be, of those around him. The writer well remembers the privilege of taking him round his own laboratory. No experiment was too unimportant for him not to wish to get a full understanding of it, which he gained with remarkable swiftness. Moreover, for every single student he had a kind and appropriate word of encouragement.

In the later years of his life a good deal of Lorentz's time and energy was occupied in delivering lectures and addresses in foreign countries and attending international scientific gatherings. Apart from his great achievements and high reputation, his peculiar powers made him an ideal president of such gatherings. He possessed and successfully employed the mental vivacity which is necessary to follow the interplay of discussion, the insight which is required to extract those statements which illuminate the real difficulties and the wisdom to lead the discussion along fruitful channels, and he did this so skilfully that the process was hardly perceptible. His linguistic powers were such that it was immaterial to him whether he spoke or listened in Dutch, English, French or German. He was adept in making singularly happy and often quite literary speeches in any one of those languages.

The writer was fortunate in being present at both the two international scientific gatherings which Lorentz attended shortly before his death, namely, the celebration at Como of the centenary of the death of Volta and the meeting of the Solvay Physics Conference at Brussels. His powers showed no perceptible diminution at either of these functions. At both he was distinguished by the alacrity of his appearance, almost that of youth, by his readiness, acuteness, and sound judgment in the debates, in which there seemed to be no subject which did not excite his interest, as well as by his kindness and friendliness with all who were present, both young and old.

For many years Lorentz had been president of the Solvay Conferences, at which a selected body of representatives of various nations meet in Brussels every three years to discuss some subject of outstanding importance in Physics. This position provides a particularly severe test on account not only of the linguistic difficulty but also of the intricacy of some of the subjects under discussion. It was remarked, however, with great satisfaction by the members present that the President accomplished his task as well as, if not better than, ever.

Lorentz received most of the honours which come to a man of the highest scientific achievements. He was elected a Foreign Member of the Royal Society in 1905 and was awarded the Rumford Medal in 1908 and the Copley Medal in 1918. He was very appreciative of his connection with the Society and delighted in recalling the close connection and intercourse between the natural philosophers of the Netherlands and England which existed in the early days of the Society. He was well known and very welcome in England. He attended the meetings of the British Association on several occasions, and, as recently as in 1923 he lectured at the Royal Institution and gave a number of lectures under the auspices of the University of London and the Anglo Batavian Society In the same year he received an honorary degree and delivered the Rede Lecture at Cambridge taking as his subject Clerk Maxwell s Electromagnetic Theory and in 1925 he gave an address to The Institute of Metals on The Motion of Electricity in Metallic Bodies

Lorentz died in the full vigour of his faculties his years had produced no perceptible waning of the wonderful spirit His death was sudden and unexpected after a short and happily painless illness he remained in good and genial spirits up to the last moment

On July 15th 1881 he married Alletta Catharina Kaiser His wife took a great interest in his various activities and was the valued companion who accompanied him on all his journeys He is survived by his widow and three children one of whom is married to Prof de Haas of Leiden and is known as the authoress of several papers on physics

The funeral took place at Haarlem at noon on Friday February 10 At the stroke of twelve the State telegraph and telephone services of Holland were suspended for three minutes as a revered tribute to the greatest man Holland has produced in our time It was attended by many colleagues and distinguished physicists from foreign countries The President Sir Ernest Rutherford represented the Royal Society and made an appreciative oration by the graveside

O W R

THEODORE W. RICHARDS.

(Listening to a critic—1911)

THEODORE WILLIAM RICHARDS—1868-1928

THIS seat is meant for three said Richards it looks small but it is safer to sit close packed humans like molecules are squeezable without breaking So Richards his wife and his guest adapted themselves to the exiguous seat at the back of the Bar Harbour Buckboard on which they proceeded to explore Mount Desert that summer home of American University presidents and professors One did not guess at the moment how naturally the analogy of a squeezable molecule came to his lips though more than once during the course of that drive a trivial incident showed the quick reaction of his thoughts On a sandy hill road they came on some youngsters drawing long lines across the sand Those are Prof Rowland s children said Mrs Richards Look how the *ruling* passion is still coming out strong said her husband

That warm September night they were rowing outside Seal Harbour in the Labrador current that cools the shores of Mount Desert and wondering at the intense phosphorescence excited by the oars Don t you find this a bit of Elysium after Harvard in July ? asked his guest Yes this is peaceful said Richards but your Oxford has its Elysium at her own doorstep I want to lie in a punt under the willows on the Cherwell that is how I dream of England Will you take me there sometime ? Sometime afterwards (it was fourteen years) the guest had the privilege of doing so

Theodore William Richards was born in Germantown Pennsylvania on January 31 1868 His parents had just returned from a residence in Germany and nearly suffered shipwreck on the homeward voyage Possibly the nervous temperament he never outgrew may be attributed in large part to this pre natal experience His father a well known painter was of English and Dutch descent his mother Anna Matlock the daughter of a Germantown physician and a member of the Society of Friends was known as a writer of religious sonnets Troubled with his eye sight Richards was not sent to school but in his mother he had a devoted and clever tutor while drawing and painting came naturally to him in that artist s home At an impressionable age he spent two years in England and he never lost his love of the English countryside

As a boy of thirteen he received a small box of chemicals and this so excited him that within a year he had fitted up a small laboratory for experiments By the kindly interest of a family friend he was invited to hear some of the Professor s chemical lectures in the University of Pennsylvania Before he was fifteen he entered Haverford College and began the serious study of Chemistry under Prof Lyman B Hall So rapidly did he progress in chemistry and physics that he came out head of his class in 1885 and graduated as B.Sc. Then, after a summer devoted to classics under his mother's tuition, he entered the senior class of Chemistry at Harvard under Prof. J. P. Cooke and resolved to make physical chemistry his life's work. In the following year he obtained the Harvard B.A. *summâ cum laude*, and then began under Prof. Cooke those investigations on the atomic weights to which he gave so much of his life and by which he established a world-wide reputation.

On completing his first research, and gaining a University fellowship and the Ph.D., he spent a session in Germany studying under Victor Meyer at Gottingen, Krüss at Munich, and Hempel at Dresden Then after a visit to England, he returned to Harvard where he became in turn Assistant (1889), Instructor (1891) and Assistant-Professor (1894). The advance of physical chemistry as a distinct branch of the Science led Richards a year later to make another visit to Germany, where he worked in Ostwald's laboratory and then with Nernst. For seven years Richards carried out the heavy teaching duties of his chair, giving all his spare time to the research laboratory—except for the summer vacation which he spent sometimes yachting at Newport, sometimes in Mount Desert, sometimes in the Adirondacks

The heavy work of the terms told on his strength, so that he was tempted by an offer of a research-professorship in Germany; but Harvard capped this offer by making him research professor, and this decided Richards to remain in his own University. This offer was not the only one made from Europe; later on he had an invitation to an important Chair in England, but he let it be known that he would only hold such a position in Europe for a limited period He might pay us a longer visit than he had made to Berlin in 1907 (when as an "Exchange Professor" he gave demonstrations on quantitative analysis to many graduates), but he felt his permanent place was on his own side of the Atlantic Indeed, his country made many calls on his time and energy, he was one of the early members of the Advisory Board of the Carnegie Institutions, he gave a series of "Lowell" lectures on the Atomic Theory, and the Willard Gibbs lecture on Atomic Weights at Chicago; he was elected Chairman of the Division of Chemistry at Harvard, President of the American Chemical Society for 1914, President of the American Association for the advancement of Science, and President of the American Academy of Arts and Sciences.

Richards began his research work on a crucial problem in chemical philosophy which greatly interested his chief—are the atomic weights of the elements multiples of that of hydrogen? In particular is Dumas' classical figure 15·96 the real value for oxygen, or did errors creep in, as Dumas himself thought, to lower the observed weight below the true whole number 16?

It was a fortunate alliance that brought the skilled hand and eye of the ardent young experimenter to carry through the scheme planned by the older chemist of revising this fundamental ratio. Prof. J. P. Cooke had shown conclusively that the high atomic weight assigned to antimony from the determinations of Dumas, of Dexter and of Kessler (over 122) was due to a constant error, and that the true value was close to 120. He agreed with Mallet that it could not be a matter of chance that so many atomic weights—compared with $O = 16$—were found to be whole numbers within the limits of experimental error. And if the majority of the well-determined atomic weights were really whole numbers, must not the variation of hydrogen from the theoretical value be apparent only? Cooke set himself to ferret out the hidden error if possible: his failing eyesight delayed the work until he was joined by Richards.

Dumas had already pointed out the weakness inherent in his process of determining the composition of water—that the hydrogen was not measured directly, but calculated by subtracting the oxygen used from the weight of water formed. A slight error in the estimation of the water, or of the oxygen, was multiplied in the hydrogen determination, and both these errors might combine to produce a great error in the hydrogen result. It was necessary, therefore, in a new determination, to weigh the hydrogen directly, and this was effected by weighing the hydrogen in a glass globe compensated for air-displacement by a similar globe on Regnault's principle When hydrogen was generated by the action of dilute sulphuric acid on zinc, Cooke and Richards found the gas contaminated by sulphur dioxide, which could only be removed by prolonged contact with potash. They therefore used hydrochloric acid, taking extreme precautions to remove dissolved air from it. With hydrogen so prepared they formed the atomic weight of oxygen almost identical with that of Dumas. They next prepared hydrogen electrolytically with zinc amalgam and hydrochloric acid, and lastly, by dissolving pure aluminium in aqueous potash. The mean results obtained with each of these three methods were wonderfully concordant; for taking $H = 1$, in the three series the values found for O were 15·954, 15·953 and 15·952 respectively.

These results, apparently so strongly confirmatory of Dumas' experimental value, were hardly in print when the authors were informed of Rayleigh's proof that a glass globe, when exhausted, sensibly contracted under the pressure of the air. Such shrinkage under exhaustion would cause the observed weight of the hydrogen filling the globe to be too small, and Richards found by weighing his globe under water that it was actually compressed in volume by 1·66 c.c The tare of the globe had thus to be corrected by 1·98 milligrams, and this reduced the atomic weight of oxygen from 15·953 to 15·869 ($H = 1$).

Since these experiments were made, there have been many elaborate determinations of the H to O ratio, of which those of Morley are among the most conclusive; but the general mean deduced from the best results does not exceed that of Cooke and Richards by 1 part in 2,000.

Richards' first independent atomic weight determination was that of copper. The accepted value 63·3 had been derived chiefly from the loss of weight of cupric oxide on reduction; Richards showed that copper oxide formed from the basic nitrate always held some occluded gas He found that pure silver was precipitated by pure copper from a solution of silver nitrate kept well-cooled. The ratio he found, confirmed by his analyses of copper sulphate, gave the accepted value of the atomic weight, 63·57 Our knowledge of the relative weights of oxygen and chlorine and thence of the atomic weights of silver, of the alkali metals and others, rested for many years on the decomposition of potassium chlorate into oxygen and potassium chloride. From the silver necessary to combine with this chloride the atomic weight of silver was derived, and from the composition of silver chloride that of chlorine itself. The results obtained by Stas were long regarded as settling these relationships with the utmost fidelity; but the "magnificent accuracy" of his work was not proof against the searching criticisms of Richards. Account had to be taken of the occlusion of oxygen by molten silver, of the action of acids on glass, of the difficulty in preventing moisture from the air being absorbed by dried substances on their way from the drying vessel to the balance—"a most frequent and insidious" source of error Simple but effective devices were invented—such as the combined "drying and bottling" tube, and the "Nephelometer" for measuring minute traces of suspended precipitates. Quartz tubes and vessels were usually employed Armed with these instruments Richards and his colleagues first determined the composition of silver chloride, and from this value proceeded to determine the atomic weights of the metals of the alkalies and of the alkaline earths by the weight of the insoluble silver chloride formed from their soluble chlorides. With similar precautions the composition of silver-bromide and that of silver iodide were determined at Harvard

In these researches Richards, with his colleagues, had established the ratio in which chlorine combines with silver and with a number of other metals, he had not, so far, brought his precise methods to bear directly on the oxygen-chlorine ratio This was accomplished by his determining the loss of weight in the reduction of pure lithium perchlorate, and the weight of silver required to precipitate the chloride remaining The atomic weight of chlorine so obtained by Richards and Willard—35·454—only differs by four parts in 35,000 from that calculated by F. W. Clarke from all the best determinations.

But the weight of an atom did not represent to Richards a more fundamental property than its volume. While gases approach a simple relation in their molecular volumes, great irregularities occur in the liquid and solid states. On applying van der Waals' equation to several gases, Richards found that the quantity *b* was not really constant, but subject to alteration under changes of temperature and pressure; but *b* represents the space actually occupied by the molecules, and if this is changeable the volume of the molecule must also be. And if this can be observed in gases, how much more evidence should be available from the expansion and contraction of solids and liquids? And not only should the force of cohesion, *e g*, in metals, result in compression, but in compounds the greater forces of chemical affinity should squeeze the atoms into smaller bulk. It is true that Davy had pointed out how much greater the contraction was when oxgyen united with potassium than when it united with tin, and other chemists had made similar observations But Richards said that the only sound basis for a generalisation lay in determining the individual compressibilities of the elements to be combined, and in measuring accurately their heats of combination.

To obtain accurate data on which to rest his hypothesis, Richards devised a simple compression vessel in which pure mercury was first compressed and the change of volume measured, and then part of the mercury being replaced by the substance to be studied, the volume change on compression gave, by difference, the compressibility of the substance--independently of any change in the vessel itself. On comparing the compressibilities of a number of elements and simple compounds, it was at once evident that the formation of a compound of a compressible element was attended with a greater decrease in volume than the formation of a similar compound of a less compressible element. Again, in determining heats of reaction, Richards overcame the difficulty of the cooling correction by the device of placing his calorimeter, like a submarine, in a vessel of alkali which he kept warming up, *pari passu* with the calorimeter, by dropping acid into it. This method proved very valuable in the study of slow reactions, and gave us, for the first time, exact data on heats of neutralisation and evaporation. With similar care, Richards studied specific heats at low temperatures—proving, indeed, that the specific heats of the metals tended to become a vanishing quantity at absolute zero. He also determined the surface tension of many liquids. The precise measurements he obtained for the compressibilities of the elements and for the heats of formation of their compounds afforded strong support for his view that the atoms themselves were close-packed and were compressible , and calculating from the data so obtained he was able to establish within approximate limits, the internal pressure of many solids and liquids due to chemical affinity and to cohesion.

However simple and reasonable Richards' view appears to us now, he had much prejudice and long opposition to overcome from those brought up in the creed of rigid and impenetrable atoms. The modern physical conception of an atom, with its outer electron-orbits far removed from the nucleus, is in agreement with Richards' hypothesis ; and he, in his turn, gave an exact confirmation of the prediction that lead derived by disintegration from thorium should differ in atomic weight but not in chemical properties from lead derived from uranium. After Richards' analyses there could be no doubt of the existence of two isotopic leads.

The photograph reproduced has the faults of the amateur's snapshot; but probably shows the real man better than more professional poses. It is from a group taken in an English garden when Richards was listening with his quiet smile before demolishing a colleague's criticism.

Richards united in his armoury two qualities not often allied in one investigator, the microscopic eye that could trace each detail, however small, and the imagination that could picture the *rerum primordia*—the hidden forces controlling all.

His life was indeed full and happy; working at full pressure of hand and brain, he yet found opportunity for making close friendships and for entertaining, and being entertained by, his scientific colleagues He found himself quite at home in England, especially in the country; and when he stayed with his friends he was not deterred from laying bare his scientific soul through any fear of being accused of "talking shop" In Germany also he was *persona grata* and had intimate friends, but when the war came he had no hesitation in expressing his sympathy with the Allies, as his letters written in the early years of the struggle show.

To the near friends who knew the sensitiveness of his nerves, and his difficulty in securing restful sleep, the work he accomplished and inspired has always been a matter of wonder and admiration Those friends will cherish the memory not only of a great man, but of a very endearing personality.

In 1896, Richards married Miriam Stuart Thayer, daughter of Dr. J. H. Thayer, the New Testament Scholar, and he leaves a daughter and two sons, The daughter Grace (Mrs J B. Conant) is married to a chemist, his elder son graduated Ph.D. in Chemistry at Harvard.

It would be useless to attempt to enumerate the honours accorded to Richards in his own and other countries; it must suffice to recall that in 1908 he was chosen an honorary fellow of the Chemical Society, before which he delivered the Faraday Lecture in 1911, he was elected a foreign member of the Royal Society in 1919, and received the Davy Medal in 1910; in 1914 he was awarded the Nobel Prize in Chemistry

Richards died comparatively young, but he had the satisfaction of knowing that his Atomic Weights, and the methods of the Harvard School he founded, had won universal acceptance; he saw his "compressible" atom no longer regarded as the fantasy of an enthusiast but as the common working-hypothesis of physicists, chemists and astronomers.

H. B. D

John I. Thornycroft.

(*By permission of the Institution of Naval Architects.*)

SIR JOHN ISAAC THORNYCROFT- 1843-1928.

The death of Sir John Isaac Thornycroft, in his 85th year, at his residence, Bembridge, Isle of Wight, means the loss of a great figure in the world of Naval Architecture and Engineering. He was responsible for many advances in the hull design of ships and in the development of machinery; and although his experiments were originally concerned with the smaller craft, his ideas exercised considerable influence upon the design and equipment of larger types of vessels.

Sir John Thornycroft was born in Rome, in which city his parents had taken up residence to study Art in the Galleries of the Vatican. (Both his father and mother were sculptors of distinction.)

Mr. Thomas Thornycroft, father of the distinguished son, was a clever amateur engineer, and, without doubt, it was due to his teaching that the boy developed his love of mechanics and obtained a considerable knowledge of engineering. It is probable, too, that Sir John's wonderful capacity for draughtsmanship was traceable in no small measure to his parents' influence.

While still a boy, he constructed his first steam launch; employing the principle of a closed stokehold and forced draught, with the propeller shaft set at a downward inclination towards the stern, permitting an exceptionally large propeller to be fitted; and the little vessel showed itself capable of developing a greater speed than other craft of similar dimensions. Shortly after this performance, the father installed his son in some small works at Chiswick, and the success of two or three launches built there was sufficient to suggest that the acquirement of further knowledge and training in this direction would afford scope for the development of a useful career. About this period he spent some time at the South Kensington School of Naval Architecture—where he was a contemporary student with Sir Philip Watts and other great naval architects. From there he went to Glasgow University, where he had the advantage of studying under Lord Kelvin and Prof. MacQuorn Rankine. He also worked for some time in George Elder's drawing office at the Fairfield Works at the time that firm was engaged in the construction of the Popofga Circular Ships for Russia. It is interesting to note that the first paper Sir John contributed to the Institution of Naval Architects (in 1869) dealt with the work he had carried out in calculating the resistance of these unusual vessels.

Returning to the Chiswick Yard, Sir John Thornycroft constructed a steam launch which showed a definite advance upon previously built craft of this class. The hull was constructed of steel—an innovation in a vessel of her size- and with an engine of about 58 h.p., a speed of 16·4 knots was obtained·

Sir Frederick Bramwell, the leading Consulting Engineer of the day, witnessed trials of this boat, and contributed a paper thereon to the Institution of Naval Architects. This performance was quickly followed by the building of another vessel, the "Gitana," which attained a speed of 23·9 m.p.h. with engines of 450 h.p.

As a consequence of the results realised with this vessel, Sir John turned his attention to the construction of torpedo boats, and in 1875 the British Admiralty placed with him the order for the "Lightning," the first torpedo boat of the British Navy. She was fitted with the closed stokehold system of forced draught, and attained a speed of 18 knots Continuing experiments with hull form and propeller design, succeeding vessels showed continuous and rapid progress. With regard to hull form, it was he who introduced the flat, wide form of stern at the water line, with the propeller shaft at downward inclination and wing rudders on each side of the stern In propeller work, he was responsible for developing turbine-screw propellers and the tunnel stern for shallow draught vessels

Sir John Thornycroft saw that the desired high speeds could not be obtained with the locomotive type of boiler hitherto employed, and took up the question of installing water-tube boilers in vessels calling for the greatest speeds Unsuccessful efforts had previously been made by marine engineers to design a boiler of this type which would give satisfactory results. In 1893, the new boilers of Thornycroft's design were given a trial in a gunboat, the "Speedy," and were entirely successful It was found by comparison with gunboats embodying the locomotive boilers, that the "Speedy" developed 4,700 h.p as against 3,500 h.p. of her sister ships, with a consequent increase in speed Ultimately Sir John produced a light, fast-running reciprocating engine of the triple expansion type, which was adopted on a large number of torpedo craft; but the reciprocating engine eventually ceased to be installed where high speeds were essential, and was superseded by the Parsons turbine. It is interesting to recall that Sir John Thornycroft was one of the two civilian members on the Fisher Committee which recommended the adoption of the Parsons turbine for the "Dreadnought" and new destroyers.

With the increased size of destroyers and torpedo boats, the Thornycroft firm moved to Southampton, and the first vessel built there, fitted with oil fuel and turbine machinery, was the fastest of the ships of this new type.

During the European War, Sir John Thornycroft was indefatigable in his work with the experimental tank, and in pushing on the construction and repair work of more than 500 ships to the order of the British Admiralty. The results of his efforts in connection therewith proved of the greatest benefit to the country in that period. It is worthy of record that throughout his long life, Sir John ever showed himself a very public spirited man in publishing information and results of his experiments.

Despite his varied work in connection with marine engineering, Sir John yet found time to conduct experiments with road motor vehicles, in their early days. These experiments were carried out in a shed at Chiswick which had formerly been used by his father as a statuary studio; and it was here that the manufacture of self-propelled carriages was later undertaken.

Sir John maintained his full vigour until a late age, and to the last continued to show the greatest interest in the work of the firm he had so ably founded, and the development of the various classes of ships, in the construction of which he had claimed so large a share during his lifetime. The firm is being carried on by his son, Sir John E. Thornycroft, who was trained as an engineer, and acted as a technical director under the father's leadership

Sir John Isaac Thornycroft became an Associate Member of the Institution of Civil Engineers in 1873, and was transferred to class of Member in 1877 He was elected a Member of the Institution of Mechanical Engineers in 1876 In 1881 he joined the Institution of Naval Architects, of which he was made Member of Council in 1882, Vice President in 1889, and an Honorary Vice-President in 1908. He was made a Fellow of the Royal Society in 1893, and received the honorary degree of LL D at Glasgow University in 1901 He was knighted in 1902.

E H T d'E.

INDEX TO VOL. CXXI. (A)

Aerofoils of small thickness (Jeffreys), 22
Aldehyde, catalytic decomposition (Allen and Hinshelwood), 141
Allen (P. C.) and Hinshelwood (C. N.) The Catalytic Decomposition of Gaseous Acetaldehyde, 141.
Alpha particles, rate of emission from radium (Braddick and Cave), 367
Alston (N. A.) and West (J.) The Structure of Topaz, 358
Ammonium salts, piezo-electric moduli (Mandell), 122 130.
Ampere meter for alternating currents (Moullin) 41
Archer (C. T.) *See* Gregory and Archer

Barlow (H. S.) *See* Williams and others
Bartlett (R. S.) The Increase of Thermionic Currents from Tungsten in Strong Electric Fields, 456
Bengough (G. D.), Stuart (J. M.) and Lee (A. R.) The Theory of Metallic Corrosion, II, 88.
Boys (C. V.) Solid Dipleidoscope Prisms—Supplement, 1
Braddick (H. J. J.) and Cave (H. M.) The Rate of Emission of Alpha Particles from Radium, 367.
Brindley (G. W.) *See* James and Brindley

Cadmium crystals, magnetic susceptibility (McLennan and others), 9
Carbon monoxide, conductivity (Gregory and Archer) 285
Carter (F. W.) On the Stability of Running of Locomotives 585
Cave (H. M.) *See* Braddick and Cave
Chalklin (F. C.) *See* Richardson and Chalklin
Chaplin (R.) The Sorption of Carbon Tetrachloride at Low Pressures by Activated Charcoals, I, 344
Cohen (E.) *See* McLennan and others
Coode Adams (W. R. C.) The Refractive Index of Quartz 176
Copper crystals, conductivity (Kannuluik and Laby), 640
Corrosion, metallic, II (Bengough, Stuart and Lee), 88
Crystal boundary, lattice spacing (Lennard-Jones and Dent) 217
Crystals, metal, banded structures (Elam), 237
Curtis (W. E.) and Harvey (A.) The Structure of the Band Spectrum of Helium 381

Dean (W. R.) Fluid Motion in a Curved Channel, 402
Dee (A. A.) *See* Smith, Dee and Young.
Dent (B. M.) *See* Lennard-Jones and Dent.
Dipleidoscope prisms (Boys), 1.
Doodson (A. T.) The Analysis and Prediction of Tidal Currents from Observations of Times of Slack Water, 72.

Eddington (A. S.) A Symmetrical Treatment of the Wave Equation, 524

Elam (C. F.) An Investigation of Some Banded Structures in Metal Crystals, 237. Appendix, by G. I. Taylor, 244.

Electrons from conductors in electric fields (Waterman), 28.

Electrons, secondary, production and absorption (Rudberg), 421.

Fisher (R. A.) The General Sampling Distribution of the Multiple Correlation Coefficient, 654.

Flint (H. T.) The New Metric of Einstein and the Wave Equation, 676.

Flow in compressible fluids (Taylor and Sharman), 194

Fluid motion in a curved channel (Dean), 402.

Fluids, compressible, flow in (Taylor and Sharman), 194.

Fluorescence and photosensitization in aqueous solutions (West, Müller and Jette), 294, 299, 313.

Freeman (L. J.) The Spectrum of Doubly-Ionised Nitrogen, 318.

Gamma-rays, internal conversion (Swirles), 447.

Goldstein (S.) The Influence of the Earth's Magnetic Field on Electric Transmission in the Upper Atmosphere, 260.

Gregory (H.) and Archer (C. T.) The Thermal Conductivities of Carbon Monoxide and Nitrous Oxide, 285.

Grindley (G. C.) *See* Tyndall and others

Harvey (A.) *See* Curtis and Harvey.

Havelock (T. H.) The Wave Pattern of a Doublet in a Stream, 515.

Helium, band spectrum (Curtis and Harvey), 381

Hinshelwood (C. N.) *See* Allen and Hinshelwood

Image force, effect on electrons (Nordheim), 626.

Ions in air, mobility (Tyndall and others), 155, 172.

Jackson (D. A.) Hyperfine Structure in the Arc Spectrum of Cæsium and Nuclear Rotation, 432

James (R. W.) and Brindley (G. W.) Reflexion of X-Rays by Sylvine, 155.

Jeffreys (H.) On Aerofoils of Small Thickness, 22.

Jette (E.) *See* West, Muller and Jette

Kannuluik (W. G.) and Laby (T. H.) The Thermal and Electrical Conductivity of Copper Crystals, 640.

Klein (F.), obituary notice of, i.

Laby (T. H.) *See* Kannuluik and Laby.

Lee (A. R.) *See* Bengough and others.

Lennard-Jones (J. E.) and Dent (B. M.) The Changes in Lattice Spacing at a Crystal Boundary, 247.

Locomotives, stability of running (Carter), 585.

Lorentz (H. A.), obituary notice of, xx.

McLennan (J. C.), Ruedy (R.) and Cohen (E.) The Magnetic Susceptibility of Single Crystals of Zinc and Cadmium, 9.

Magnetic susceptibility of zinc and cadmium crystals (McLennan and others), 9.

Mandell (W.) The Change in Elastic Properties on Replacing K of Rochelle Salt by Ammonium, 122. The Determination of the Piezo-Electric Moduli of Ammonium Seignette Salt, 130.

Moullin (E. B.) An Ampere Meter for Alternating Currents of very High Frequency, 41.
Müller (R. H.) *See* West, Müller and Jette.

Neumann bands, mode of formation (Smith and others), 477, 486, 501.
Nitrogen, the spectrum of doubly-ionised (Freeman), 318.
Nitrous oxide, conductivity (Gregory and Archer), 285.
Nordheim (L. W.) The Effect of the Image Force on the Emission and Reflexion of Electrons by Metals, 626.
Nuclear rotation, hyperfine structure (Jackson), 432.
Nuttall (J. M.) *See* Williams and others.

Obituary notices.—
Klein (F.), i.
Lorentz (H. A.), xx
Richards (T. W.), xxix.
Thornycroft (Sir John), xxxv.

Particles, random velocities from encounters (Thomas), 464.
Photosensitization in aqueous solutions (West, Müller and Jette), 294, 299, 313
Piezo-electric properties of certain salts (Mandell), 122, 130.
Powell (C. F.) *See* Tyndall and others.

Quartz, refractive index (Coode Adams), 476

Richards (T. W.), obituary notice of, xxix.
Richardson (O. W.) and Chalklin (F. C.) The Soft X-Ray Levels of Iron, Cobalt, Nickel and Copper, 218.
Robertson, (A.) The Strength of Tubular Struts, 558.
Rudberg (E.) The Production and Absorption of Soft X-Rays and Secondary Electrons, 421.
Ruedy (R.) *See* McLennan and others.

Sharman (C. F.) *See* Taylor and Sharman.
Sheppard (P. A.) *See* Tyndall and others.
Smith (S. W. J.), Dee (A. A.) and Young (J.) The Mode of Formation of Neumann Bands, 477, 486, 501.
Sorption of carbon tetrachloride at low pressures (Chaplin), 344.
Spectrum, arc, of cæsium (Jackson), 432.
Spectrum, band, of helium (Curtis and Harvey), 381.
Spectrum of doubly-ionised nitrogen (Freeman), 318.
Starr (L. H.) *See* Tyndall and others.
Stuart (J. M.) *See* Bengough and others.
Swirles (B.) The Internal Conversion of Gamma-Rays, 447
Sylvine, reflection of X-rays (James and Brindley), 155.

Taylor (G. I.) Appendix on Banded Structures in Crystals. *See* Elam.
Taylor (G. I.) and Sharman (C. F.) A Mechanical Method for solving Problems of Flow in Compressible Fluids, 194.
Temple (G.) The Scattering Power of a Bare Nucleus according to Wave Mechanics, 673.
Thermionic currents from tungsten in strong electric fields (Bartlett), 456.
Thomas (L. H.) On the Rate at which Particles take up Random Velocities from Encounters according to the Inverse Square Law, 464.
Thornycroft (Sir John), obituary notice of, xxxv.

Tidal currents, analysis and prediction (Doodson), 72

Topaz, structure (Alston and West), 358.

Tubular struts, strength (Robertson), 558

Tyndall (A M), Grindley (G C) and Sheppard (P A) The Mobility of Ions in Air, V, 185

Tyndall (A M), Starr (L H) and Powell (C F.) The Mobility of Ions in Air, IV, 172.

Upper atmosphere, earth's magnetic field and electric transmission (Goldstein), 260

Waterman (A T) The Effect of Electric Fields on the Emission of Electrons from Conductors, 28

Wave equation and the new metric (Flint), 676

Wave equation, symmetrical treatment (Eddington), 524

Wave mechanics, principle of least action (Whittaker), 543

Wave mechanics, scattering power of a bare nucleus (Temple), 673

Wave pattern of a doublet in a stream (Havelock), 515

West (J) *See* Alston and West

West (W), Müller (R H) and Jette (E) Studies on Fluorescence and Photosensitization in Aqueous Solution, 294, 299, 313

Whittaker (J M) On the Principle of Least Action in Wave Mechanics, 543

Williams (E J), Nuttall (J M) and Barlow (H S) The Special Distribution of Photo Electrons produced by X Rays, 611

X ray levels, *soft, of metals* (Richardson and Chalklin), 218

X rays, soft, production and absorption (Rudberg), 421

X rays, photo electrons produced (Williams, Nuttall and Barlow), 611

Young (J) *See* Smith, Dee and Young

Zinc crystals magnetic susceptibility (McLennan and others) 9

END OF THE ONE HUNDREDTH AND TWENTY FIRST VOLUME (SERIES A)

R. I. 75.

IMPERIAL AGRICULTURAL RESEARCH
INSTITUTE LIBRARY
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.
12.5.47		
5.12.47		
26.11		
7.8.51		